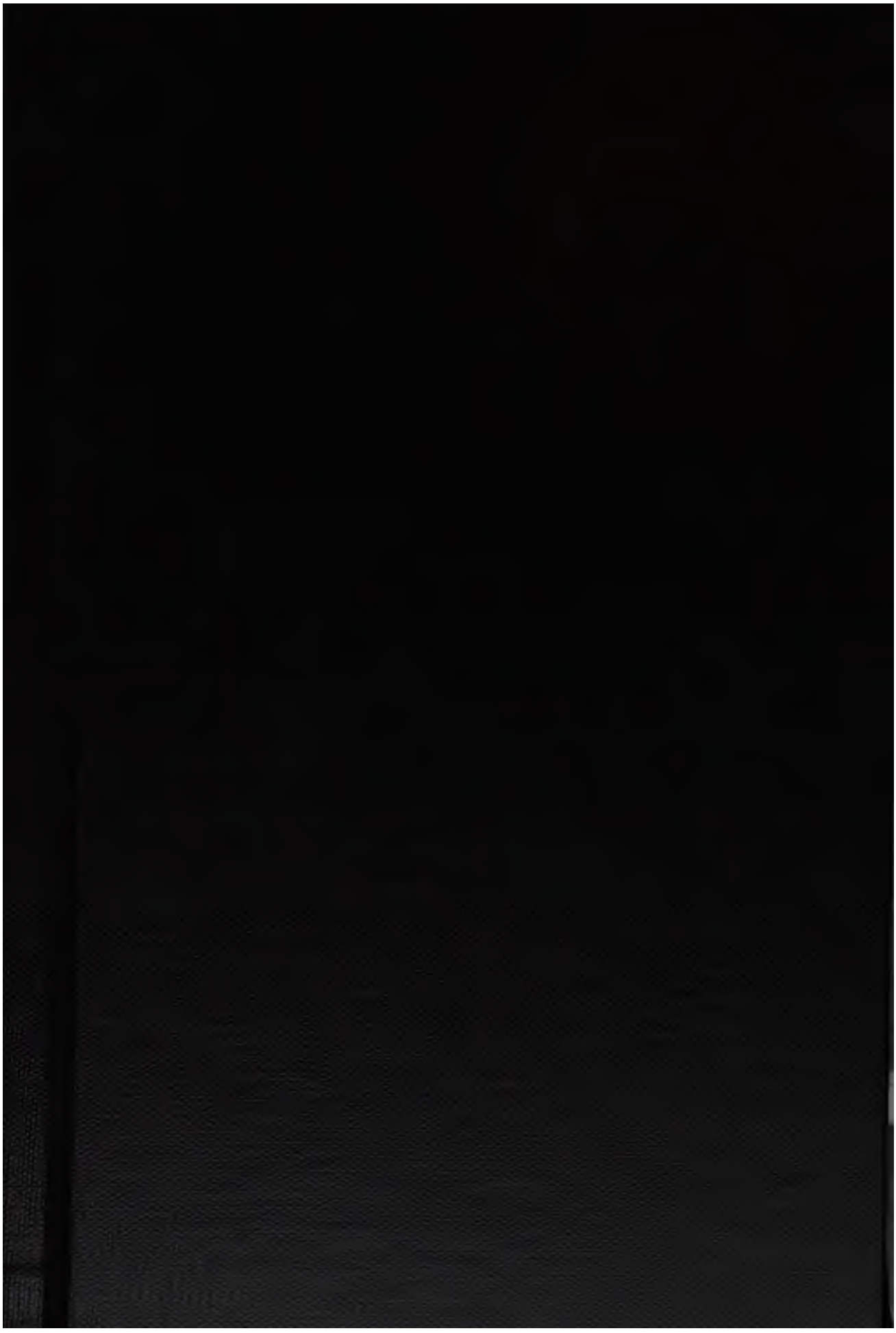



STANFORD UNIVERSITY
LIBRARIES


STANFORD UNIVERSITY
LIBRARIES

Whole No. 1 Series 1, No. 1
HARVARD MONOGRAPHS IN EDUCATION

A COMPARISON OF THE INTELLIGENCE AND TRAINING OF SCHOOL CHILDREN IN A MASSACHUSETTS TOWN

BY
EDWIN A. SHAW
AND
EDWARD A. LINCOLN
PSYCHO-EDUCATIONAL CLINIC, HARVARD UNIVERSITY

Series 1 No. 1

STUDIES IN EDUCATIONAL PSYCHOLOGY
AND
EDUCATIONAL MEASUREMENT

Edited by
WALTER F. DEARBORN

MAY, 1922

Published by
THE GRADUATE SCHOOL OF EDUCATION
HARVARD UNIVERSITY, CAMBRIDGE, MASS.

INTRODUCTION

Most of the previously published School Surveys have dealt with the schools of a rather large community. This fact has undoubtedly led those interested in the schools of the smaller towns to feel that surveys were not practicable or valuable for any but the large systems with thousands of pupils, hundreds of teachers, a considerable investment in school property, and plenty of money to meet the expenses of such investigations. This report is published in the hope that it will show the possibilities of the survey on a small scale. The school "system" with which the report deals consisted of only two schools, with scarcely more than 200 pupils, and less than 20 teachers. In spite of these small numbers and in spite of the fact that the survey was necessarily limited in its scope, it is the conviction of practically everyone concerned that the results have been of considerable value in improving the efficiency of the schools. The cost was little more than a dollar a pupil.

It is only fair to state that the community derived great advantages from the fact that the testing was done, papers were corrected, and the results were worked up as laboratory assignments of classes in Educational Measurement under the direction of the writers. This relieved the teachers of all work, and also materially cut down the financial outlay necessary. However, it must be held in mind that the work could have been done by the regular teachers under the direction of the superintendent or some other leader who had given some time to the systematic study of Educational Measurement under a competent instructor. In such circumstances the cost would be even lower than it was in the present instance.

Just as Part I of this report was ready for the press there came an invitation to repeat the survey. Some reorganization had taken place as the result of the findings which had been reported, and the School Board was desirous of learning the result of these changes. It was considered worth while to delay this report pending the preparation of Part II, in order that the value of repeated measurements might be shown. The delay seems to have been justified, for Part II shows not only some changes which have come about as a result of the first survey, but also points out some new work which needs to be done in the schools.

The two reports are printed exactly as they were sent to the School Board, with the exception that the class lists containing each

pupil's record in the tests are omitted. A sample of them is shown in the appendix. The discussion, of course, has been made as non-technical as possible, since the reports were intended primarily for the lay reader. The expert in educational measurement or the executive in a large school system will perhaps find little to interest him in this report. For those, however, who have felt that the paucity of pupils and the stringency of finances must deprive the small school system of the advantages of educational measurement we trust it may prove an inspiration and a guide.

The authors wish to express their deep appreciation of the earnest and painstaking labor of their students in the preliminary work on which this report is based. They also acknowledge their indebtedness to Dr. Walter F. Dearborn for his helpful criticism in every phase of their undertaking.

EDWIN A. SHAW.
EDWARD A. LINCOLN.

PSYCHO-EDUCATIONAL CLINIC.
HARVARD UNIVERSITY.

REPORT ON THE RESULTS OF TESTING IN THE WAYLAND AND COCHITUATE SCHOOLS

PART I—THE SURVEY OF 1921

PURPOSE OF THE REPORT

It is the purpose of any school survey, whatever may be its nature and its scope, to discover and place at the disposal of all those vitally interested in any phase of the educational activities of the community definite facts concerning the work which is going on in the schools. To censure or to praise is not a primary function of the survey; its first and most important object must be to determine with absolute impartiality and scientific accuracy the existing conditions in the school or school system. Constructive criticism by disinterested outsiders is, of course, valuable and necessary, but teachers and administrative officers, because they know in greater detail their schools and the community are, theoretically, at least, better able to make final decisions concerning the activities of the schools. This report, then, will place before the teachers, superintendent, school board, and citizens of the town certain facts about the pupils and the work of the schools with which they are so vitally concerned. Such comment and criticism will be made as, in the judgment of the surveyors, will contribute to the improvement of the educational opportunity offered at present by these schools.

HOW THE SCHOOL WORK WAS MEASURED

All modern school surveys make use of the standard tests in school subjects. As will be seen from the descriptions which appear later in the report, the standard tests are not in any great measure different from the ordinary school examinations which are more or less familiar to everyone who has ever attended school. The difference arises from the fact that the standard tests are scientifically constructed after long study and experimentation, and they have been given to many thousands of children all over the country, so that the average or standard performance in a room, grade, school, or whole system is pretty well known. Thus it is possible by the use of standard tests to compare the work of the various units of a school system, and also to compare the work done in the system with that done in any other where the same

tests have been used. For those reasons standard tests were used in the survey of the Wayland and Cochituate schools.

WHAT TESTS WERE USED

Although standard tests have been devised in a large number of the subjects commonly taught in the schools, not all of them could be used in this survey. Because of the limitations, tests were given in only the fundamental subjects; namely, Reading, Writing and Arithmetic. To these was added an intelligence test for the purpose of determining to some extent the character of the raw materials with which the various units of the school system have to work. Obviously, more is to be expected of a class composed of pupils of superior intelligence than from the average class, and inferior pupils cannot be expected to do as well as the average.

HOW THE TESTS WERE GIVEN

All the testing was done by examiners who had been well trained for the work at the Psycho-Educational Clinic, of the Harvard Graduate School of Education. Although the standard tests are for the most part so devised that they can be given by the average class-room teacher, there is a technique of testing which is acquired only after considerable practice in the field, and it is, therefore, always desirable to have survey work done by trained examiners, as in the present instance.

CORRECTION OF PAPERS

The papers were all corrected by the clerks at the Psycho-Educational Clinic. Here, also, much of the work can be done by the teacher, but it is more efficiently and economically done by trained experts. Moreover, when comparisons are to be made, it is an advantage to have the work done as far as possible by disinterested and impartial outsiders.

PRESENTATION OF RESULTS

All the results are presented in tabular form. The first set of tables presents the results of the tests by grade and school, and each table shows a distribution of all the marks obtained by the pupils of each grade, together with the grade median score. The Median rather than the Average is used because it is much more easily found, and because most of the standard results are expressed as median scores.

The median is simply the middle score when all the scores are arranged in order of their size.*

A second set of tables will show the Wayland and Cochituate medians together with those from other systems with which comparisons are desired, and in the Appendix will be found a sample of the class lists which will indicate the score of each pupil in each of the tests which he took.

TABLE 1. DEARBORN GROUP INTELLIGENCE EXAMINATIONS

General Examination I, May, 1921

Distribution of Scores by Grade†

	Cochituate				Wayland		
	Sub-Primary	I	II	III	I	II	III
120+				4			
110-119				2			
100-109		1	2	3			
90-99		2	2	1		2	
80-89		0	3	1		1	
70-79	1	1	4	2	1	4	
60-69	1	3	4	1	0	1	Given
50-59	2	2	1	0	2	4	Series
40-49	4	3	0	0	1	2	II
30-39	6	4	0	1	2		
20-29	12	3	1		1		
10-19	7		1		0		
0-9					1		
No. Cases	33	19	18	15	8	14	
Median	28	48	75	105	43	70	

*Thus the median of 5, 6, 8, 11, 12, 15, 20 is 11. When there is an even number of items in the series the median is the same as it would be if one more item were added to the upper end of the series. Of the series 5, 6, 8, 10, 12, 14, 18, 19, 20, 21, the median is 14.

†For the lay reader who is not accustomed to reading tables of this sort it may be said that figures in the above table indicate the number of pupils who got each score on the examination. Thus in the Cochituate First grade 3 pupils made scores between 20 and 29, inclusive; 4 pupils made scores between 30 and 39, inclusive; 3 scored between 40 and 49, inclusive, and so on.

TABLE 2. DEARBORN GROUP INTELLIGENCE EXAMINATIONS

General Examination I, May, 1921

Distribution of Scores by Age

	Cochituate					Wayland		
Years	5	6	7	8	9	6	7	8
Score								
120+			1	4	1			
110-119			0	2	0			
100-109		1	0	0	3		1	
90-99		0	1	2	2		0	1
80-89	1	0	1	2	0		0	1
70-79	0	0	2	1	3	1	2	2
60-69	0	2	2	3	2	0	0	0
50-59	0	3	1	1	0	2	2	2
40-49	1	2	2	2	0	1	2	
30-39	0	5	2	3	0	1	1	
20-29	8	4	2		2		1	
10-19	2	3	1				0	
0-9							1	
No. Cases	12	20	15	20	13	5	10	6
Median	25	36	55	80	78	53	50	75

RESULTS OF INTELLIGENCE EXAMINATIONS

The Dearborn Group Intelligence Examinations which were used in this survey are made in two forms, one for the younger pupils who have not yet mastered the arts of reading and writing well enough to use them to any great practical extent, and the other for the older boys and girls. In the present instance the complete lower grade examination was not used, but a shorter form, which has also been standardized, was employed. This examination will be referred to in this report as General Examination I.

A distribution of the scores made on General Examination I in the different grades of both schools is presented in Table 1. Because of the grouping of the grades it was found more practicable to test the third grade in Wayland with the Series II examination, so no results for this grade are to be found in Table 1.

It will be seen that the pupils in the Cochituate school are, in both the first and second grades, somewhat superior to the pupils in the same grades at Wayland. We may reasonably expect, therefore, when we come to the subject-matter tests, to find the median scores of the first two grades in Cochituate somewhat higher than those of the corresponding grades in Wayland.

In Table 2 the distribution of the General Examination I scores is according to the age of the pupils. The range in age is from 5 to 9 years in Cochituate, and from 6 to 8 only in Wayland. This difference is accounted for by the fact that in Cochituate there were Sub-Primary and Third grades tested, while at Wayland the test was given to the First and Second grades alone.

TABLE 8. DEARBORN GROUP INTELLIGENCE EXAMINATIONS

Series II, May, 1921

Distribution of Scores by Grade

	Cochituate					Wayland						
Grades	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII	H.S.
200+												
190-199					1							
180-189					0							2
170-179					0						1	5
160-169					2						0	3
150-159				1	0						0	9
140-149				1	1						0	3
130-139		1		2	1						0	3
120-129	1	0	4	1	4						1	5
110-119	0	1	1	4	2			1	1		0	5
100-109	0	0	1	2	2			1	0	4	3	2
90-99	2	2	1	2	0		1	1	1	3	1	2
80-89	2	1	1	0	3		0	0	0	2	0	1
70-79	7	3	1	2		1	2	1	2	0	0	1
60-69	2	1	0			1	1	3	1	1	0	
50-59	7	2	2			0	0	1	1	1	1	
40-49	5	2	0			3	5	1	0			
30-39	2	2	1			5	2	0	2			
20-29						1	2	1	2			
10-19						1						
0-9												
No. Cases	28	15	12	15	16	12	13	10	10	11	7	41
Median	50	72	100	114	123	38	45	67	60	95	105	145

TABLE 4. DEARBORN GROUP INTELLIGENCE EXAMINATIONS
Series II, May, 1921
Distribution of Scores by Age

	Cochituate							Wayland							
Age	9	10	11	12	13	14	15	9	10	11	12	13	14	15	16
Score															
200+															
190-199						1							2		
180-189						0							0		
170-179						0						1	0	1	3
160-169						0						1	2	1	0
150-159				1		1	1					0	0	0	0
140-149				0	1	1						0	0	1	2
130-139			1	0	0	3						2	1	0	0
120-129		1	1	0	4	3	1					1	0	0	2
110-119		1	0	2	3	1	1			1	1	0	0	2	1
100-109		0	0	1	2	1	1		1	1	2	0	1	2	1
90-99	2	1	1	1	0	2		1	1	0	2	2	1	0	
80-89	0	1	2	0	0	1	2	0	0	0	0	1	1	1	
70-79	3	2	2	3	1	1		2	2	1	1	0	0		
60-69	2	1	0	0	0			0	3	0	2	0	1		
50-59	5	1	0	3	2			0	0	2	1	1	0		
40-49	5	2	1	1				5	3		1	0	0		
30-39	1	2	2					2	1		0	0	1		
20-29								2	1		1	1	1		
10-19								1							
0-9															
No. Cases	16	12	10	12	13	15	6	13	12	5	11	10	11	8	9
Median	54	70	80	77	115	135	110	43	63	75	75	110	105	115	143

It will be seen that the six-year-old children in Wayland make higher scores than the Cochituate pupils of the same age, but that for the seven and eight-year-olds the Cochituate medians are better. It must be noted that there are so few six and eight year old pupils in Wayland that the medians for these years are rather unreliable.* However, the weight of the evidence points again to the fact that the Cochituate pupils are somewhat superior.

Tables 3 and 4 show the distributions of the scores made on the Series II examinations which were given to the older pupils. Here again we find evidence of superiority of the Cochituate pupils, espe-

*The reliability of a median is proportional to the number of cases from which it is obtained. See Thorndike, Mental and Social Measurements, Chap. XII, Reliability of Measures.

cially in the three upper grades. A noteworthy fact is the low standing of the sixth grade in Wayland. The median for this grade is only 60, as compared with a median of 67 in the next lower grade. Other discrepancies of the same sort appear when the scores are distributed according to age in Table 4. The median score of the twelve year olds is no greater than that of the eleven year olds, and the thirteen year olds score, on the average, higher than the fourteen year olds. In the fifteenth year the Wayland median is 5 points higher than that of Cochituate. This is explainable by the fact that in Cochituate the fifteen year olds are probably for the most part backward pupils who have not yet managed to get out of the grades, while the fifteen year olds from Wayland include the brighter pupils of that age who have gone on to High School.

TABLE 6: PEET-DEARBORN ARITHMETIC TESTS

Addition

May, 1921

	Cochituate						Wayland				
Grade	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
Score											
130											
125-129											
120-124											
115-119		1			1						
110-114	1	1									
105-109		1									
100-104	1			1	2			1			
95-99					1						
90-94		1									
85-89				1	1						
80-84				1							
75-79	1		1		2						
70-74			1	1			1			1	
65-69				3	1			1			
60-64			1	1							4
55-59		1		1	2			1	2		3
50-54		1	1		1				2		1
45-49	2	1			3			2			
40-44	3		1	2				3	1	2	
35-39	4	2	2	1					1		2
30-34	2	4	1	2			4		1	3	
25-29	2				1	1	4			1	
20-24	3			1	1	1	1	1	1	2	
15-19	6		2			2	1	1	1	1	
10-14			2			1	1				
5-9	2					6	3			1	
0-4	1					1	2		1	2	
Total	28	13	12	15	16	12	17	10	10	13	10
Median	30	43	38	63	60	9	26	45	40	28	58
Standard	30	51	35	51	54	..	30	51	34	51	54

TABLE 7. PEET-DEARBORN ARITHMETIC TESTS

Subtraction

May, 1921

	Cochituate						Wayland				
Grade	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
Score											
130		1		1				1			
125-129											
120-124					1						
115-119		1			1			1			
110-114											
105-109					1		1				
100-104				1							
95-99			1	1	3						
90-94		1		1							
85-89	1		1	2	2		1	1			
80-84				1							
75-79		1		1	1		1	1			3
70-74		1		1	1					1	
65-69		1	1		2					1	1
60-64	2	2	1	3						1	1
55-59	1						1	1			
50-54		2	1					1	1	2	1
45-49	1		1							1	2
40-44	5			3			2		1		
35-39	4							2		1	
30-34	2		5		4		4	1	1	.	1
25-29	1						3			1	
20-24	2	1				4	1	1			1
15-19	3	1	1			2	2		2	2	
10-14	1						1		2	2	
5-9	1					4			2	1	
0-4	4	1				2			1		
Total	28	13	12	15	16	12	17	10	10	13	10
Median	35	64	40	78	80	13	32	55	15	38	55
Standard	29	51	34	52	62	..	29	51	34	52	62

TABLE 8. PEET-DEARBORN ARITHMETIC TESTS

Multiplication

May, 1921

	Cochituate						Wayland				
Grade	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
Score											
130				1	2			1			
125-129											
120-124		1									
115-119		2			1						
110-114					1						
105-109				1							
100-104					2						
95-99		1		1	3						1
90-94	1	3		1				2		1	
85-89		1	1	1	2			2			
80-84			1	1							
75-79		1		1	1						
70-74			1								1
65-69				1			1				
60-64	1			4				1	1	1	1
55-59		1	1	2						1	2
50-54	1		1	1	2		1	1			
45-59	1	1			2			1			1
40-44	1		1				4	1		1	1
35-39	1		1				1			1	2
30-34	2		1				3		4	4	
25-29	1		2				1		1		1
20-24	4	1	2				2		2	1	
15-19	3					5	2		1		
10-14	3						2	1			
5-9	2					6					
0-4	7	1				1			1	3	
Total	28	13	12	15	16	12	17	10	10	13	10
Median	18	91	40	63	97	9	33	75	31	43	55
Standard	28	48	41	52	58	..	28	48	41	52	58

TABLE 9. PEET-DEARBORN ARITHMETIC TESTS

Division

May, 1921

	Cochituate						Wayland				
Grade	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
Score											
130					5						
125-129											
120-124											
115-119				2				1			
110-114					2						
105-109								1			
100-104											
95-99		1									
90-94											
85-89		2	1	2							
80-84											
75-79	1	1		4	1			2			1
70-74		3	3		1			2			2
65-69			1	1	1						
60-64	1	1	2	3	1					1	2
55-59				1					1		
50-54	2		1	1				1	1		1
45-49			2	1			1		1	2	
40-44	5	1			1				1	3	2
35-39	2		1							1	1
30-34	5				1		1	1	1		
25-29	1	1			3					1	
20-24	3		1			3	4		1	3	1
15-19	3	1				2	6	2			
10-14	3	1				1			2		
5-9	1					5	3		2	1	
0-4	1	1				1	2		1	1	
Total	28	13	12	15	16	12	17	10	11	13	10
Median	32	71	63	76	75	10	18	73	18	38	53
Standard	28	49	31	49	61	..	28	49	31	49	61

The facts concerning the Addition examples are shown in Table 6. Here the discrepancies between the corresponding grades of the two schools are not great, except in the Seventh grade scores. In this grade the Cochituate median is 35 points higher than the Wayland median.

Table 7 shows the results of the Subtraction test. Here again there are variations in favor of Cochituate which range from 6 to 27 points.

In the Multiplication test (Table 8) the Fourth grade in Cochituate did very much poorer work than it did on the previous tests, and the Wayland median is higher in this grade. In all the other grades, however, Cochituate did considerably superior work, with medians ranging from 9 to 27 points higher than those of the Wayland classes.

The test in Division (Table 9) shows the Cochituate work superior except in the Fifth grade. It is also worthy of note that the Eighth grade in Cochituate does not attain a higher median than the Seventh grade. It is, however, considerably above the standard median.

In summary of the results of the Arithmetic tests it may be said that the work in Cochituate is considerably in advance of that at Wayland, with only a few exceptions. The difference amounts, on the average, to about a year's work. This is the state of affairs which was predicted upon examination of the results of the intelligence tests, and it is not, therefore, surprising.

TABLE 10. PICTURE SUPPLEMENT SILENT READING TEST

Distribution of Scores

May, 1921

	Cochituate						Wayland					
Grade	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
100												
98												
92												
86												1
80	1											0
74	1	1										0
68	0	4		1		1						0
62	1	3		0	1	0			1			0
56	2	4	1	1	0	0		1	2		1	0
50	3	5	1	2	2	1	1	0	0		0	1
44	1	3	0	1	1	3	1	2	1		2	2
38	2	1	1	0	0	1	2	2	3	2	0	2
32	4	1	1	0	1	4	1	4	1	0	0	1
26	3	1	2	3	5	4	5	2	0	1	1	1
20	1		7	1	2	3	0	1	1	2	5	0
14			1	0	1		0	1	0	0	3	0
8				0			0	1	1	1	0	1
2				1			0	0		0	1	
0							1	2		2		
No. Cases	19	23	14	10	13	17	11	16	10	8	13	9
Median	38	56	20	25	26	32	26	32	38	20	20	38

RESULTS OF THE READING TEST

The test used in this survey for the measurement of reading is one recently published by the Russell Sage Foundation, and is known as Picture Supplement Scale, Form 1. It consists, according to the author's description, of "a series of pictures and paragraphs about them. These paragraphs consist of instructions which the pupil follows by marking with his pencil a line or lines to supplement the picture. His ability to do this in accordance with printed instructions reflects the rapidity and accuracy with which he can read." A pupil's score is the number of paragraphs he marks correctly, but in order that all the grades may be put on the same basis the author provided a table of credits corresponding to the number of paragraphs correctly marked in each grade.

The distribution of the scores on the credit basis is given in Table 10, together with the median credit attained in each grade. For the most part, the Cochituate median credits are the higher, but in the Eighth grade Wayland has the slight advantage of one more paragraph correctly marked. This result again accords with the findings of the intelligence examinations which led us to expect better work from the pupils in Cochituate.

RESULTS OF THE PENMANSHIP TESTS

The Holmes Test for the Speed and Quality of writing was used in this investigation. For the speed test, a short sentence, made up of simple and familiar words, but involving nearly all the letters and a large number of the common letter combinations, is written repeatedly, first for a one minute period, and then for four minutes. The sentence is practically learned by the children before they begin to write on the test, so that little time is lost because of forgotten copy. The speed of writing is obtained by taking the average of the number of letters written in the one-minute period and the number written per minute in the four minute period.

TABLE 11. HOLMES TEST, SPEED OF WRITING. May, 1921

	Cochituate						Wayland					
Grade	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
Letters per min.												
120+						5					1	1
115-119					2	1					0	1
110-114					3	4					0	1
105-109			1		1	2					0	0
100-104			1	1	1	2					1	1
95-99			0	2	2	1					1	1
90-94			0	1	1	1				1	2	3
85-89	1		2	1	2				1	0	2	1
80-84	0		1	2	1				1	1	2	
75-79	0	1	3	0	0			1	1	0	1	
70-74	2	5	3	0	0			3	0	0	1	
65-69	0	5	0	2	1			0	4	2		
60-64	2	3	1	0				2	2	2		
55-59	2	5	1	0			2	0	0	1		
50-54	2	1	0	0			3	3	0	0		
45-49	1	5	1	1			1	3	1	1		
40-44	4	1					2	1		1		
35-39	0	0					1	2				
30-34	0	1					1					
25-29	1						0					
20-24							1					
19-												
No. Cases	15	27	14	10	14	16	11	15	10	9	11	9
Median	60	61	77	85	100	113	48	53	68	63	88	98

TABLE 12. HOLMES TEST, QUALITY OF WRITING. May, 1921

	Cochituate							Wayland					
Grade	II	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
90-94													
85-89													
80-84													1
75-79											1	1	
70-74							2			3		1	1
65-69							2						
60-64						3	1			2	1	1	1
55-59			2		1	2				1	1		3
50-54		3	2	1	1	3	7		1	3		2	1
45-49		2	8	5	3	3	3		2		4	2	2
40-44	3	4	9	3	4	2	1	2	5			2	
35-39	1	4	1	3	1	1		3	3	1	1		
30-34	6	2	1	2				4	4			1	
25-29	4	4						2	1				
20-24	2												
15-19													
10-14													
Total	16	19	23	14	10	14	16	11	16	10	8	10	9
Median	32	39	46	43	45	50	53	32	40	60	48	50	57

The facts concerning the speed of writing are shown in Table 11. Here again the medians show that the work at Wayland is about a grade below that at Cochituate. The low standing of the Sixth grade at Wayland is once more noticeable.

To get the quality of penmanship by the Holmes test, two samples of each child's writing are graded on the Ayres Handwriting scale. The first of these samples is the last half of the material written in the four minute speed test, and for the other the pupils are given a short passage to write from dictation. The samples are graded separately, each by two correctors working independently, and the average of the four marks (two on each paper) is taken as the quality of the pupil's handwriting.

The distributions and medians of the marks obtained in the quality of penmanship are shown in Table 12. In this test we find for the first time that the Wayland medians equal or surpass those of Cochituate. The difference in most cases is small, amounting to about half a step on the Ayres Scale, but in the Fifth grade the penmanship at Wayland is especially good.

TABLE 11. HOLMES TEST, SPEED OF WRITING. May, 1921

	Cochituate						Wayland					
Grade	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
Letters per min.												
120+						5					1	1
115-119					2	1					0	1
110-114					3	4					0	1
105-109			1		1	2					0	0
100-104			1	1	1	2					1	1
95-99			0	2	2	1					1	1
90-94			0	1	1	1				1	2	3
85-89	1		2	1	2				1	0	2	1
80-84	0		1	2	1				1	1	2	
75-79	0	1	3	0	0			1	1	0	1	
70-74	2	5	3	0	0			3	0	0	1	
65-69	0	5	0	2	1			0	4	2		
60-64	2	3	1	0				2	2	2		
55-59	2	5	1	0			2	0	0	1		
50-54	2	1	0	0			3	3	0	0		
45-49	1	5	1	1			1	3	1	1		
40-44	4	1					2	1		1		
35-39	0	0					1	2				
30-34	0	1					1					
25-29	1						0					
20-24							1					
19-												
No. Cases	15	27	14	10	14	16	11	15	10	9	11	9
Median	60	61	77	85	100	113	48	53	68	63	88	98

TABLE 12. HOLMES TEST, QUALITY OF WRITING. May, 1921

	Cochituate							Wayland					
Grade	II	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
90-94													
85-89													
80-84													1
75-79											1	1	
70-74							2			3		1	1
65-69							2						
60-64						3	1			2	1	1	1
55-59			2		1	2				1	1		3
50-54		3	2	1	1	3	7		1	3		2	1
45-49		2	8	5	3	3	3		2		4	2	2
40-44	3	4	9	3	4	2	1	2	5			2	
35-39	1	4	1	3	1	1		3	3	1	1		
30-34	6	2	1	2				4	4			1	
25-29	4	4						2	1				
20-24	2												
15-19													
10-14													
Total	16	19	23	14	10	14	16	11	16	10	8	10	9
Median	32	39	46	43	45	50	53	32	40	60	48	50	57

The facts concerning the speed of writing are shown in Table 11. Here again the medians show that the work at Wayland is about a grade below that at Cochituate. The low standing of the Sixth grade at Wayland is once more noticeable.

To get the quality of penmanship by the Holmes test, two samples of each child's writing are graded on the Ayres Handwriting scale. The first of these samples is the last half of the material written in the four minute speed test, and for the other the pupils are given a short passage to write from dictation. The samples are graded separately, each by two correctors working independently, and the average of the four marks (two on each paper) is taken as the quality of the pupil's handwriting.

The distributions and medians of the marks obtained in the quality of penmanship are shown in Table 12. In this test we find for the first time that the Wayland medians equal or surpass those of Cochituate. The difference in most cases is small, amounting to about half a step on the Ayres Scale, but in the Fifth grade the penmanship at Wayland is especially good.

COMPARISONS WITH OTHER SYSTEMS

1. The Arithmetic Tests

In the following tables, 13 to 17 inclusive, are shown the Cochituate and Wayland median Arithmetic scores compared with the standards.

TABLE 13. PROBLEMS
May, 1921

Grade	IV	V	VI	VII	VIII
Cochituate	18	63	28	56	73
Wayland	14	35	5	11	40
Standard	24	43	22	38	47

In the Fourth grade neither school attains the standard median score for the Problem Test, but in the Fifth grade the Cochituate median is considerably above the standard, while the Wayland median is considerably lower. In the upper grade test which was given in the grades above the Fifth, the Cochituate medians are uniformly considerably larger than the standards, and the Wayland medians are considerably smaller.

TABLE 14. ADDITION TEST
May, 1921

Grade	IV	V	VI	VII	VIII
Cochituate	30	43	38	63	60
Wayland	26	45	30	28	58
Standard	30	51	35	51	54

In the Fourth grade, the Cochituate median in Addition exactly equals the standard, but the Wayland score falls below. Neither school equals the Fifth grade standard. In the Sixth and Seventh grades the Cochituate medians surpass the standards, while the Wayland scores fall below again, especially in the Seventh grade. In the Eighth grade, both scores exceed the standard by a good margin.

TABLE 15. SUBTRACTION TEST
May, 1921

Grade	IV	V	VI	VII	VIII
Cochituate	35	64	40	78	80
Wayland	32	55	15	38	55
Standard	29	51	34	52	62

In grades Four and Five the Subtraction medians in both schools are higher than the standards. In the upper grade test we find the Cochituate medians exceeding the standards by a good margin, especially in the Seventh and Eighth grades. Wayland in the upper grades falls below the standard, and has an especially low score in grade Six.

TABLE 16. MULTIPLICATION TEST

May, 1921

Grade	IV	V	VI	VII	VIII
Cochituate	18	91	40	63	97
Wayland	33	75	31	43	55
Standard	28	48	41	52	58

In multiplication the Cochituate Grade Four median falls below the standard, but Wayland is above it. In the Fifth grade both schools exceed the standard by large variations. The upper Wayland grades cannot attain the standards in this operation, but the Eighth grade comes very close to it. The Cochituate median is practically the same as the standard in the Sixth grade, and in the Seventh and Eighth grades the medians of this school are especially high.

TABLE 17. DIVISION TEST

May, 1921

Grade	IV	V	VI	VII	VIII
Cochituate	32	71	63	76	75
Wayland	18	73	18	38	53
Standard	38	49	31	49	61

In the Fourth grade neither school attains the Division standard, though the Cochituate median is very close. Both schools are well above the standard in the Fifth grade. In the three upper grades we again find that the Cochituate medians are well above the standards for those grades, while the Wayland medians fall below.

TABLE 18. PICTURE SUPPLEMENT, SILENT READING TEST

Form I, May, 1921

Per Cent of Pupils Receiving Each Mark

	Cochituate							Wayland					
Grade	III	IV	V	VI	VII	VIII	Standard	III	IV	V	VI	VII	VIII
100													
98							1—						
92							1+						
86							3						
80	5						4						
74	5	4		10			6						
68	0	18		0		6	8						
62	5	13		10	8	0	10			10			
56	11	18	7	0	0	0	11		6	20		8	
50	16	22	7	20	15	6	12	9	0	0		0	
44	5	13	0	10	8	18	11	9	13	10		15	11
38	11	4	7	0	0	6	10	18	13	30	25	0	22
32	21	4	7	0	8	23	8	9	25	10	0	0	22
26	16	4	15	30	39	23	6	46	13	0	13	8	11
20	5		50	10	15	18	4	0	6	10	25	38	0
14			7	0	8		3	0	6	0	0	23	0
8				0			1+	0	6	10	13	0	11
2				10			1—	0	0		0	3	
0								9	13		25		
No. Cases	19	23	14	10	13	17		11	16	10	8	13	9

2. The Reading Test

Table 18 sets forth a comparison of the scores in the Silent Reading Test with the standards. The scoring of this test is so arranged that in any average grade from the Second to the Eighth inclusive the percentage of pupils obtaining each score is the same. It will be readily seen by reference to the Table that in neither school does the distribution equal or approximate the standard. There is no grade in either school which has the average quota in the upper third of the distribution.

3. The Penmanship Tests

In Table 19 are shown the medians for the speed of writing. Together with the standards there are in this Table the scores from three Massachusetts communities where the Holmes Test has been given.

TABLE 19. SPEED OF PENMANSHIP

May, 1921

Grade	III	IV	V	VI	VII	VIII
Cochituate	60	61	77	85	100	113
Wayland	48	53	68	63	88	98
Standard	49	62	76	87	90	98
Newton	55	59	73	85	94	102
Brookline			76	87	90	98
Fall River				86	89	

The Cochituate medians surpass the standards by a good margin in every grade except the Fifth, and also exceed the medians of the other school systems which are listed in the Table. Wayland, on the other hand, shows median scores considerably below the standards of the other school systems up until the Seventh and Eighth grades, where the standards are equalled.

The medians for the quality of penmanship are shown in Table 20. Here we find that, in general, the medians are higher than the standards, and higher than those obtained in Brookline and Fall River. They do not quite equal the Newton scores in the grades below the Seventh

TABLE 20. QUALITY OF PENMANSHIP

May, 1921

Grade	III	IV	V	VI	VII	VIII
Cochituate	39	46	43	45	50	53
Wayland	32	40	60	48	50	57
Standard	36	39	44	46	47	49
Newton	50	45	48	51	50	53
Brookline			44	46	47	49
Fall River				44	47	

SUMMARY

Our comparison of the two schools shows, in general, with the exception of a grade or two here and there, that the quality of work done by the Cochituate pupils is decidedly superior to that done at Wayland. This statement holds except for the quality of penmanship, in which Wayland does especially well. An explanation of the better performances of the Cochituate pupils is found in the fact that they make higher scores on the Intelligence Examinations, thus indicating the probability that they are somewhat superior in their native abilities.

It is possible also to make a rather general statement about the comparisons of the Cochituate and Wayland scores with the standards. For the most part, the Cochituate medians are equal to, or better than the

standards, while the Wayland scores equal the standards in only a few cases out of the total. Again there is an exception to this in the case of the quality of penmanship, in which test the Wayland pupils equalled or surpassed the standards in most of the grades.

SUGGESTIONS

It is a fact sometimes overlooked in a report of this sort that interpretation of the results set forth is chiefly the business of the teachers and supervisory officials of the school system in which the tests have been given. No one else is continually on the ground and familiar with all the conditions; no one else can understand the facts in their full significance.

Perhaps the most valuable use to which standard tests may be put by a school system is that of serving as a guide for the formulation of standards for its own work. Tests are first given to show the actual attainment of the pupils in the schools. With these results and information concerning the results of testing in other schools standards may be set in the various school subjects.

It must be held in mind continually, when such standards are being determined, that they must always be made with reference to the conditions in the system in question. Each community has its own individuality—its own problems, its own potentialities, and its own limitations. The recent development of intelligence tests has demonstrated that there are undoubtedly differences in native capacity which appear to exist between schools and school systems as well as between individuals. No system, therefore, may safely take as its own, standards which have been made elsewhere, but each must work out a set of standards for itself.

In this connection it should be pointed out that there has been somewhat of a tendency to set too high the standards in the school subject tests. It is forgotten sometimes that these tests, for the most part, measure only the simpler and more mechanical phases of a subject, and that excessive attention to these phases may mean that more important things are neglected. Thus it becomes a question, for instance, whether the relatively high scores made by the Cochituate pupils on most of the tests show a desirable state of affairs, and whether the school should aim to keep future classes at the same level. If the superior scores are due to superior ability, as may be the case, it is to be desired that the same high grade of work be continued, but if these scores are due to over-emphasis on the mechanical parts of the various subjects, then some change should be made.

This brings us to another important consideration: namely, that the standards should always be tentative. It is impossible to determine once and for all what the accomplishments of pupils should be. From

time to time, as the community changes, the standards should change also. In other words, the problem of standards is always an open and live one.

Once tentative standards have been set, the tests are useful to determine the progress of the various schools in relation to those standards, and to gather material for the revision of them. For these purposes the tests should be repeated at not too frequent intervals; once a year is probably often enough.

The results should be studied with the purpose of picking out schools, classes and individual pupils to which some sort of special attention should be given. One school of a system may do especially well in writing, but fall down in reading, while another may exceed the standards in both reading and writing, but make a very poor showing in the arithmetic. A class may stand high in addition, but fail in division. Some pupils will be found to be in need of much individual work, while others will seem to have attained satisfactory proficiency in all lines. In the lower part of the group there will be found cases which should be turned over to a specialist for searching mental examinations, while there will be others whose attainments will warrant double promotion, or who should be enabled to speed up their work in some other way.

Superintendents and teachers should remember that neither every individual nor every class should be expected to attain the standards. If anything stands out clearly in testing, it is the fact of individual differences, and many pupils will never be able to reach the standard medians. Indeed, the very definition of the term "median" indicates that it is a mark which only one-half of the pupils equal or exceed.

It is often the practice of supervising officials to judge the worth of their teachers by the scores that the pupils in the various classes make on the standard tests. It is more likely, however, that the success which a class has in making high scores in the standard tests has little or nothing to do with the ability of the teacher. Since the assignment of teachers to classes is largely a matter of chance, it may easily happen that the best class in the school may have the poorest teacher, and that the success of a class may be in spite of the teaching, rather than because of it.

To sum up, then, if the facts set forth in this report are to become useful they must be studied carefully and exhaustively by the people interested in the schools. It must be remembered that high scores in the tests are not the only, nor even the most desirable ends of school work, and that possibly the classes which do poorly in these are getting fully as

much from the school. The results in the subject tests must be considered in relation to the intelligence scores, and to all other facts concerning the general conditions in the schools and in the community. A little follow-up work, with special attention along the lines suggested above can hardly fail to be of inestimable value to the pupil, the teacher, and the society of which the pupil is to become a part.

PART II—THE SURVEY OF 1922

In May, 1922, a year after the first tests had been given in Wayland and Cochituate the whole survey was repeated. The tests used were the same, except that there had been some slight modifications of the Intelligence and Arithmetic Tests. The testing and grading was again done by carefully trained examiners.

The results of this second testing are set forth in tables which are numbered to correspond with the tables in Part I, so that it will be easy for anyone who is interested to compare the distributions for the two years.

RESULTS OF INTELLIGENCE EXAMINATIONS

Distributions of the scores made by the pupils of the lower grades on the Dearborn General Examination A are shown in Table 1a. In Table 2a the same scores are distributed according to the ages of the pupils.

There is evidence in these figures that Cochituate has pupils of superior mentality in the First grade, as shown by the median score of 48 as opposed to the Wayland median of 33. Also, it may be seen in the second Table that the median score of the six-year-olds is identical in both schools. In the Second grade, however, it appears that the Wayland children are superior, with a class median of 66 against 60 made by the Cochituate class. The seven-year-olds at Wayland also do somewhat better than the seven-year-olds at Cochituate. The superiority of the Wayland children in the Second grade amounts to about five months on a mental age basis, and while this is not an especially great difference, it is probably large enough to be significant.

TABLE 1a. DEARBORN GROUP INTELLIGENCE EXAMINATIONS

General Examination A, May, 1922

Distribution of Scores by Grade

	Cochituate				Wayland		
	Sub-Primary	I	II	III	I	II	III
95-99				2			
90-94		1		3			
85-89				3			
80-84			1	4			
75-79			0	3		3	
70-74			2	1		0	
65-69			2	2		2	
60-64		2	0	3		1	Given
55-59	1	6	3	2		2	Series
50-54	0	3	1	0		0	II
45-49	0	4	1	1	1	1	
40-44	2	6	1		0		
35-39	2	2			4		
30-34	6	2			3		
25-29	7	2			1		
20-24	1	1			1		
15-19	1				0		
10-14	1				1		
5-9	1				0		
0-4					1		
No. Cases	22	29	11	24	12	9	
Median	30	48	60	80	33	66	
Mental Age of Median Child	6-8	7-7	8-4	10-5	6-10	8-9	

TABLE 2a. DEARBORN GROUP INTELLIGENCE EXAMINATIONS

General Examination A, May, 1922

Distribution of Scores by Age

	Cochituate						Wayland		
Years	5	6	7	8	9	10	6	7	8
95-99				1					
90-94				1	3	1			
85-89			2	1	0	0			
80-84			1	0	0	2			
75-79			0	0	2	1		2	1
70-74			0	1	1	1		0	0
65-69			1	2	1	0		1	1
60-64			1	1	2	0		1	0
55-59		4	2	3	2	1		1	1
50-54		0	3	1		0		0	0
45-49		1	1	3		1		1	1
40-44	1	1	4	3				0	
35-39	0	1	3	0			1	3	
30-34	2	3	2	1			2	1	
25-29	1	7	1				1		
20-24	0	2					1		
15-19	0	1							
10-14	1								1
5-9	1								
0-4							1		
No. Cases	6	20	21	18	11	7	6	10	5
Median	30	30	48	56	73	78	30	55	58
M.A. of Med. Child	6-8	6-8	7-7	8-1	9-6	10-1	6-8	8-0	8-2

NOTE:—Two 11-year old pupils and one 13-year old from the Cochituate school are omitted from this Table.

In Tables 3a and 4a will be found the distributions of the Dearborn General Examination C which was given to the older children. In the Fourth and Sixth grades the Wayland medians are higher, in the Fifth and Seventh grades the Cochituate medians are higher, while in the Eighth grade the medians are exactly the same. The largest difference is in the Sixth grade where Wayland has a median which is 8 points higher than that of Cochituate. This means a difference of over a year in terms of mental age, and is a large variation. The difference of four points in the Fourth grade is probably also large enough to be of significance.

This state of affairs is quite different from that which was discovered the previous year. In 1921 the results showed rather clearly that in most grades the pupils of the Cochituate school were superior to those in the same grade at Wayland. Now many of the differences seem to have disappeared, and those which remain seem to be in favor of Wayland rather than of Cochituate. A possible explanation of this change lies in the fact that the results of the first report were used to make extensive changes in the organization of each school, and it seems likely that the reorganization classified the children into much more homogeneous groups.

One interesting fact should be noted in passing. In 1921 the Sixth grade at Wayland proved to be inferior to the fifth grade in the intelligence test. In 1922 the Seventh grade, composed largely of the same pupils who made up the Sixth grade in the previous year, is quite markedly inferior to the Sixth. Evidently the character of this class was not entirely changed in the reorganization.

TABLE 8a. DEARBORN GROUP INTELLIGENCE EXAMINATIONS

Gen. Exam. C, May, 1922

Distribution of Scores by Grade

	Cochituate					Wayland					
Grade	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
80+											
75-79											
70-74					1						
65-69					1						
60-64			1	2	3				2	1	1
55-59			1	1	2				1	1	5
50-54		1	1	1	4				1	0	3
45-49		2	0	0	2			2	2	2	3
40-44		3	4	4	2			0	1	2	1
35-39	1	2	3	1	1			0	1	1	0
30-34	3	6	2	3	1		1	2	1	3	1
25-29	2	6	1	0		2	3	2	1		0
20-24	2	6	1	1		2	3	4	0		1
15-19	3	8	2			6	2	1	1	2	
10-14	2	1	1			2	1	1			
5-9	2					1	1				
0-4	2										
No. Cases	17	35	17	13	17	13	11	12	11	12	15
Median	19	27	38	42	53	18	23	25	46	40	53
M.A. of Med. Child	10-2	11-6	13-4	14-0	15-10	10-0	10-10	11-2	14-8	13-8	15-10

TABLE 4a. DEARBORN GROUP INTELLIGENCE EXAMINATIONS

Gen. Exam. C, May, 1922

Distribution of Scores by Age

	Cochituate							Wayland							
Age	9	10	11	12	13	14	15	8	9	10	11	12	13	14	15
80+															
75-79															
70-74					1										
65-69				1	0										
60-64				1	2	2				1	1	2			
55-59		1		1	0	2				1	1	0	1	1	3
50-54		0	1	2	1	3	1			0	0	1	0	2	0
45-49		2	0	0	1	1				2	1	1	3	2	0
40-44		2	1	5	2	3				0	0	2	2	0	0
35-39	1	0	2	2	3	0				0	0	1	1	0	0
30-34	2	4	1	3	1	4			1	1	2	1	1	1	1
25-29	1	3	3	0	1	1			4	2	2				
20-24	2	2	3	1	2			2	2	4	1				
15-19	3	4	3	3				6	0	2	1		2		1
10-14		2	1	1				1	1	1	1				
5-9		2	0						1	1					
0-4		1	1												
No. Cases	9	23	16	20	14	16	1	9	9	15	10	8	10	6	5
Med'n	24	26	25	40	40	45		18	26	24	30	45	43	50	
M.A. of Med. Child	11-0	11-4	11-2	13-8	13-8	14-6		10-0	11-4	11-0	12-0	14-6	14-2	15-4	

NOTE: One 16-year old pupil and one 18-year old pupil from the Wayland school have been omitted from this Table.

RESULTS OF THE ARITHMETIC TESTS

As previously noted, the Peet-Dearborn Arithmetic Tests had been somewhat modified between the two testings so the results of the two years are not directly comparable. There are still, however, two series of the tests, one for the Fourth and Fifth grades, and the other for grades Six, Seven and Eight. This accounts for the fact that in all the Tables the Sixth grade scores are usually lower than those made by the Fifth grade.

In Table 5a will be found the distributions of the scores obtained on the Problem Test, together with the class medians. The medians in grades IV and V show better work by the Wayland children, though the Fifth grade medians differ by only two points, an insignificant difference. In the upper grades, however, the scores are very decidedly in favor of Cochituate. These differences cannot be explained on the basis of the intelligence of the classes, except perhaps in the Fourth grade where the Wayland class made a significantly higher median in the intelligence test. It would seem that in the Sixth grade especially, where the mental age of the Wayland pupils is about a year higher, a larger median might reasonably be expected.

Table 6a presents the distributions and medians in the Addition Test. Here we find that the medians for the Fourth and Fifth grades are practically the same, though in the former grade the Cochituate median is somewhat higher. In the upper grades we find the Cochituate medians are larger by substantial margins which range from 10 to 21 points.

In Table 7a we have the facts about the Subtraction Test. It will be seen that in the three lower grades the Wayland medians are decidedly higher, while in the Seventh and Eighth grades the Cochituate classes show to much better advantage. It appears that the Seventh and Eighth grade medians at Wayland are extremely low.

TABLE 5a. PEET-DEARBORN PROGESS TESTS IN ARITHMETIC
Problems
May, 1922

	Cochituate					Wayland				
Grades	IV	V	VI	VII	VIII	IV	V	VI	VII	VIII
130+					2					
125-129										
120-124										
115-119										
110-114					1		2			
105-109					0			1		2
100-104					0					
95-99					2					
90-94					5					1
85-89			1		0					0
80-84		2	1	1	0					3
75-79		1	0	1	2					0
70-74		4	1	0	1					1
65-69		0	0	2	1				2	0
60-64		2	0	0	0	1			0	0
55-59	1	3	3	0	1	0	2		1	0
50-54	0	1	4	4	1	1	2	2	1	0
45-49	0	1	0	0		0	0	0	0	0
40-44	0	5	0	1		0	2	2	2	3
35-39	0	0	0	2		1	1	0	1	2
30-34	0	0	0	0		0	0	1	0	0
25-29	1	2	2	1		1	0	0	1	1
20-24	3	5	2	1	1	2	1	0	1	0
15-19	2	0	2			0	0	0	1	1
10-14	3	2	1			1	0	2	2	1
5-9	5	2				4	1	2		
0-4	2	3					1	1		
No. Cases	17	33	17	13	17	11	12	11	12	15
Median	13	43	52	52	91	22	50	33	40	44
Standard	24	38	22	41	56	24	38	22	41	56

TABLE 6a. PEET-DEARBORN PROGRESS TESTS IN ARITHMETIC

Addition

May, 1922

	Cochituate					Wayland				
Grades	IV	V	VI	VI	VIII	IV	V	VI	VII	VIII
130+		1								
125-129										
120-124							2			
115-119					1			1		
110-114										
105-109										
100-104		2								
95-99		0						1		1
90-94		0								0
85-89	1	4	1		4					2
80-84		1			0					0
75-79		2			1			2	2	0
70-74	1	1	1	1	2		2	0	1	1
65-69	0	2	0	0	2		1	0	1	0
60-64	1	5	2	5	3	1	4	0	0	2
55-59	0	1	2	0	1		2	1	0	0
50-54	0	0	2	2	0				1	1
45-49	0	2	2	1	2				0	1
40-44	3	1	5	0	0				1	2
35-39	1	2	1	2	1	1		2	0	0
30-34	2	6	1	2		3		1	1	1
25-29	2	1				0		0	0	0
20-24	5	2				2		1	4	3
15-19	1					3	1	2	1	
10-14						1				
5-9										
0-4										1
N.o Cases	17	33	17	13	17	11	12	11	12	15
Median	31	62	49	54	69	24	64	39	38	48
Standard	30	50	35	52	60	30	50	35	52	60

TABLE 7a. PEET-DEARBORN PROGRESS TESTS IN ARITHMETIC

Subtraction

May, 1922

	Cochituate					Wayland				
Grades	IV	V	VI	VII	VIII	IV	V	VI	VII	VIII
130+										
125-129										
120-124										
115-119										
110-114		1	1		2					
105-109		0					1			
100-104		1								
95-99								1		
90-94					1		1	0		1
85-89		3			1	1	0	1		0
80-84		1		1	0		1	1		1
75-79		1		0	3		1	0		1
70-74		4	2	2	1		1	1		1
65-69	1	0	0	0	1		1	1		1
60-64	1	3	3	3	2	2	2	1		0
55-59	1	4	0	0	0	1	1			0
50-54	2	3	3	1	2	2	0			1
45-49	0	2	1	2	0	0	0			0
40-44	3	2	2	0	3	2	1		2	1
35-39	1	1	0	1	1	1	1		1	0
30-34	1	3	1	0		1	0	1	0	3
25-29	1	1	1	0		0	0	0	3	0
20-24	5	1	1	1		1	1	3	2	2
15-19	0	1	1	1					1	3
10-14	1			1					1	
5-9									1	
0-4								1	1	
No. Cases	17	32	16	13	17	11	12	11	12	15
Median	38	58	52	53	68	51	65	63	25	34
Standards	28	47	39	57	68	28	47	39	57	66

Distributions and medians for the Multiplication Test are shown in Table 8a. The Fourth grade medians are practically identical. Wayland has an advantage of 8 points in the Fifth grade, and Cochituate excels in the upper grades by margins running up to 32 points in grade Eight.

Table 9a shows the results of the Division Test. Here we find that the Cochituate medians are decidedly higher throughout the whole five grades. The points of difference range from 9 in the Sixth grade to 38 in the Eighth.

In Table 9b the averages of the pupils in all operations are distributed. The medians are substantially the same in the Fourth and Fifth grades, but in the upper grades the Cochituate medians are considerably higher.

Summarizing the results of the Arithmetic Tests we may say, in general, that the work of the Cochituate children is superior. This is not true in all cases in the two lower grades (IV and V) where the medians are more nearly equal and where the Wayland medians are in one or two cases distinctly higher. The differences which we find are not to be explained on the basis of intelligence differences, as the latter are small except in cases where they favor Wayland. It seems likely, therefore, either that more time is given to arithmetic, especially in the upper grades, at Cochituate than at Wayland, or that the Cochituate teachers use markedly superior methods in their teaching.

TABLE 8a. PEET-DEARBORN PROGRESS TESTS IN ARITHMETIC

Multiplication

May, 1922

	Cochituate					Wayland				
Grades	IV	V	VI	VII	VIII	IV	V	VI	VII	VIII
130+					1					
125-129										
120-124										
115-119										
110-114					1		1			
105-109		2			0		0			
100-104		3			3		2	1		
95-99		1		1	0		0			
90-94		1		0	0		0			1
85-89		6		3	3		2		1	1
80-84		0		0	1	1	1		0	1
75-79		3		3	2			1	1	0
70-74	1	1	2	0	3			0		2
65-69	0	0	1	0	0			1		
60-64	0	2	5	1	2	1				
55-59	2	0	0	0	0	0	2			1
50-54	3	1	1	1	1	1	0			2
45-49	2	3	1			0		1		1
40-44	2	2	1			4	1	1	2	3
35-39	3	1	0			0	0	2	0	1
30-34	2	3	4	2		1	0	2	2	2
25-29		1	1	1		1	1	0	2	
20-24		1		0		1	1	2	2	
15-19	1	1		1		0			1	
10-14						1			1	
5-9										
0-4	1									
No. Cases	17	32	16	13	17	11	11	11	12	15
Median	44	75	60	76	83	42	83	39	30	51
Standard	28	48	31	51	68	28	48	31	51	68

TABLE 9a. PEET-DEARBORN PROGRESS TESTS IN ARITHMETIC

Division

May, 1922

	Cochituate					Wayland				
Grades	IV	V	VI	VII	VIII	IV	V	VI	VII	VIII
130+		1			1					
125-129										
120-124				1						
115-119							1			
110-114					1					
105-109		1			1					
100-104		1		1	3					1
95-99		0	1	0	0					0
90-94		1	1	1	0					1
85-89		2		2	2			1		
80-84		0		0	0			0		
75-79		3		2	4			1		
70-74		1		0	0		1	0		2
65-69		3		0	1	1	1	1	1	0
60-64	2	3	3	1	2	0	1	0	1	1
55-59	0	2	0	1		0	0	0	1	0
50-54	2	3	3	0		1	1	1	0	1
45-49	1	2	2	1		0	1	0	2	0
40-44	3	2	2	1		1	2	2	2	4
35-39	3	3	1	0	1	1	1	0	0	3
30-34	3	2	3	2	0	1	1	4	0	1
25-29					1	0	1	0	1	0
20-24						3	1	0	2	1
15-19	1					1		1	2	
10-14	0	1				1				
5-9	2	1				1				
0-4										
No. Cases	17	32	16	13	17	11	12	11	12	15
Median	39	61	50	76	79	24	45	41	43	43
Standard	29	46	34	55	67	29	46	34	55	67

TABLE 8a. PEET-DEARBORN PROGRESS TESTS IN ARITHMETIC

Multiplication

May, 1922

	Cochituate					Wayland				
Grades	IV	V	VI	VII	VIII	IV	V	VI	VII	VIII
130+					1					
125-129										
120-124										
115-119										
110-114					1		1			
105-109		2			0		0			
100-104		3			3		2	1		
95-99		1		1	0		0			
90-94		1		0	0		0			1
85-89		6		3	3		2		1	1
80-84		0		0	1	1	1		0	1
75-79		3		3	2			1	1	0
70-74	1	1	2	0	3			0		2
65-69	0	0	1	0	0			1		
60-64	0	2	5	1	2	1				
55-59	2	0	0	0	0	0	2			1
50-54	3	1	1	1	1	1	0			2
45-49	2	3	1			0		1		1
40-44	2	2	1			4	1	1	2	3
35-39	3	1	0			0	0	2	0	1
30-34	2	3	4	2		1	0	2	2	2
25-29		1	1	1		1	1	0	2	
20-24		1		0		1	1	2	2	
15-19	1	1		1		0			1	
10-14						1			1	
5-9										
0-4	1									
No. Cases	17	32	16	13	17	11	11	11	12	15
Median	44	75	60	76	83	42	83	39	30	51
Standard	28	48	31	51	68	28	48	31	51	68

TABLE 9a. PEET-DEARBORN PROGRESS TESTS IN ARITHMETIC

Division

May, 1922

	Cochituate					Wayland				
Grades	IV	V	VI	VII	VIII	IV	V	VI	VII	VIII
130+		1			1					
125-129										
120-124				1						
115-119							1			
110-114					1					
105-109		1			1					
100-104		1		1	3					1
95-99		0	1	0	0					0
90-94		1	1	1	0					1
85-89		2		2	2			1		
80-84		0		0	0			0		
75-79		3		2	4			1		
70-74		1		0	0		1	0		2
65-69		3		0	1	1	1	1	1	0
60-64	2	3	3	1	2	0	1	0	1	1
55-59	0	2	0	1		0	0	0	1	0
50-54	2	3	3	0		1	1	1	0	1
45-49	1	2	2	1		0	1	0	2	0
40-44	3	2	2	1		1	2	2	2	4
35-39	3	3	1	0	1	1	1	0	0	3
30-34	3	2	3	2	0	1	1	4	0	1
25-29					1	0	1	0	1	0
20-24						3	1	0	2	1
15-19	1					1		1	2	
10-14	0	1				1				
5-9	2	1				1				
0-4										
No. Cases	17	32	16	13	17	11	12	11	12	15
Median	39	61	50	76	79	24	45	41	43	43
Standard	29	46	34	55	67	29	46	34	55	67

TABLE 9b. PEET-DEARBORN PROGRESS TESTS IN ARITHMETIC

Average Operations

May, 1922

	Cochituate					Wayland				
Grades	IV	V	VI	VII	VIII	IV	V	VI	VII	VII
130+										
125-129										
120-124										
115-119					1					
110-114							1			
105-109		1								
100-104		0			1					
95-99		1								
90-94		1								
85-89		2	1	2	2					3
80-84		1		1	3		2	1		
75-79		2		0	2		0			
70-74		2	1	0	3		1			1
65-69		3		2	1		2	1	1	
60-64		6		2	0			1	1	
55-59	1	1	2	2	1			0		1
50-54	2	4	4	0	2	1		1		1
45-49	1	0	4	1	1	1	1	1		1
40-44	5	2	1	0		3	1	1	3	2
35-39	3	1	2	1		1	1	1	2	2
30-34	1	3	1	1		2	2	2	0	2
25-29	1	0		0		2		0	2	1
20-24	2	1		1		1		2	1	1
15-19	0	1							1	
10-14	1								1	
5-9										
0-4										
No. Cases	17	32	16	13	17	11	11	11	12	15
Medians	41	63	50	61	76	38	66	43	38	44
Standard	30	47	34	53	65	30	47	34	53	65

RESULTS OF THE READING TEST

The Picture Supplement Test for silent reading is published in several forms of equivalent value, and as Form 1 was used in 1921, Form 2 was chosen for the 1922 testing. The distributions and medians will be found in Table 10. It should be remembered in connection with these figures that the scoring of this test is arranged to give a median score of 50 to the average or normal class, irrespective of the grade. In the Third and Fourth grades the differences are not very great, and are slightly in favor of the Cochituate classes. In the Fifth and Sixth grades, however, Wayland shows decidedly to advantage. Cochituate medians are higher in the Seventh and Eighth grades.

TABLE 10a. BURGESS PICTURE SUPPLEMENT SILENT READING TEST
Distribution of Scores
May, 1922

	Cochituate						Wayland					
Grades	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
100									1	2		
98												
92						1						
86	2				1	0						
80	0		1		1	1				1	..	1
74	1		0	1	0	1				1	1	0
68	2	1	2	3	2	4			3	3	1	1
62	1	2	3	3	0	0	3		2	1	0	2
56	1	3	3	1	2	0	2	1	1		1	3
50	5	0	1	0	0	5	0	0	1		2	0
44	3	4	4	1	0	1	2	2	0		0	0
38	6	2	5	2	2	0	2	5	1	1	0	2
32	1	3	4	2	1	1	1	0	0	1	1	3
26	2		7	0	3	1	2	1	1	1	2	2
20			2	3	1			1	2		2	1
14			1								1	
8		1	1								0	
2											1	
0								1				
No. Cases	24	16	34	16	13	15	12	11	12	11	12	15
Median	50	47	40	53	43	55	47	41	62	71	32	43

These results seem to imply some differences in the methods of teaching reading or the subject matter used or both. Further, the differences seem to exist not only between the two schools, but between the different grades in each separate school.

RESULTS OF THE PENMANSHIP TEST

The Holmes Test was used, exactly as in the previous survey, and the children were rated for both Speed and Quality of writing.

The scores for Speed in letters per minute are shown in Table 11a. The Cochituate medians in every grade except the Eighth are higher than the Wayland medians. The differences, however, are hardly great enough to be of significance except in the Seventh grade.

The distributions of the scores in Quality of penmanship are shown in Table 12a. The differences in the grade medians of the two schools are so small as to be insignificant, excepting possibly the Third grade, where the difference is five points, or half a step on the Ayres scale, in favor of Wayland. In neither school is there any improvement in the Quality of the writing after the Sixth grade.

HOLMES TEST

TABLE 11a. SPEED OF WRITING

May, 1922

	Cochituate							Wayland					
Grades	II	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
125-129													
120-124													
115-119					2		1						2
110-114				1	2		1						0
105-109				1	1	1	3						2
100-104				1	2	2	2			1			4
95-99				2	6	2	2			0	2	1	1
90-94				7	2	4	0			0	4	1	4
85-89				4		2	5			2	1	0	1
80-84			1	1		2	2			2	0	6	0
75-79		..	1	4						0	2	2	1
70-74	1	1	2	3	1					2	0	3	
65-69	1	3	3	4			1			1	1		
60-64	0	1	0	2					1	0			
55-59	1	0	1	4				2	1	1			
50-54	1	4	6.					1	1	2	1		
45-49	2	8	2					1	3				
40-44	1	4	0					4	0				
35-39	3	1	1					2	1				
30-34	0	2						1	1				
25-29	1							1	1				
20-24									1				
15-19													
10-14													
5-9													
No. Cases	11	24	17	34	16	13	17	12	10	11	11	13	15
Med'n	46	48	54	80	99	94	96	42	46	74	91	82	101

HOLMES TEST

TABLE 12a. QUALITY OF WRITING

May, 1922

	Cochituate							Wayland					
Grades	II	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
95-99													
90-94													
85-89													
80-84													
75-79													
70-74													
65-69													1
60-64													1
55-59													0
50-54							2				1	3	2
45-49				3	3	1	2	2			1	0	2
40-44			3	4	3	0	3	1		3	2	2	2
35-39			3	7	4	8	4	2	2	3	3	2	0
30-34	1	3	3	7	4	3	3	2	5	1	3	5	4
25-29	6	13	6	6	2	1	1	3	1	4	1	1	2
20-24	4	8	2	7			1	2	3	1			1
15-19													
10-14													
5-9													
No. Cases	11	24	17	34	16	13	16	12	11	12	11	13	15
Med'n	26	28	32	33	38	37	39	33	31	35	38	36	41

COMPARISONS WITH STANDARDS

1. The Arithmetic Tests

In the following Tables 13a to 17a inclusive will be found the median arithmetic scores of both schools in comparison with the standards.

TABLE 13a. PROBLEMS

May, 1922

Grade	IV	V	VI	VII	VIII
Cochituate	13	43	52	52	91
Wayland	24	45	33	40	44
Standard	24	38	22	41	56

The Fourth grade at Cochituate and the Eighth grade at Wayland are the only grades which fall below the standards in the Problem Test. In the other grades the medians are pretty uniformly well above the

standards. It is worthy of note that there has been a decided improvement in the work at Wayland, for the results of the first testing showed not a single median which equalled or surpassed the standards in this test.

TABLE 14a. ADDITION TEST
May, 1922

Grade	IV	V	VI	VII	VIII
Cochituate	31	62	49	54	69
Wayland	24	64	39	38	48
Standards	30	50	35	52	60

In Addition the record is again good. The Cochituate medians are all higher than the standards, as are the Wayland Fifth and Sixth grade medians. In the Fourth grade the Wayland median is somewhat below the standard. The Seventh grade shows no progress over the Sixth, and the Eighth grade also is low.

TABLE 15a. SUBTRACTION TEST
May, 1922

Grade	IV	V	VI	VII	VIII
Cochituate	38	58	52	49	68
Wayland	51	65	63	25	34
Standards	28	47	39	57	66

In Subtraction the Cochituate medians are well above the standards except in the Seventh grade, where both schools fall below. The Wayland Eighth grade attained a median which is only about half what may be expected from the average class. The Wayland record is somewhat better than the previous one, as this time there are three instead of two classes above the standards. It is also noticeable that in the Fourth, Fifth and Sixth grades the Wayland medians are well above those of the Cochituate classes.

TABLE 16a. MULTIPLICATION TEST
May, 1922

Grade	IV	V	VI	VII	VIII
Cochituate	44	75	60	76	83
Wayland	42	83	39	30	51
Standards	28	48	31	51	68

In the Multiplication Test the Cochituate pupils obtained especially high scores, as the medians show. We again find the Wayland medians surpassing the standards in the first three grades, but they fall below in the Seventh and Eighth. The records are not greatly different from

those made at the previous testing except for the fact that the Cochituate Fourth grade and the Wayland Sixth grade have much better records in the 1922 tests.

TABLE 17a. DIVISION TEST
May, 1922

Grade	IV	V	VI	VII	VIII
Cochituate	39	61	50	76	79
Wayland	24	45	41	43	41
Standard	29	46	34	55	67

The Cochituuate class medians are all well above the standards in the Division Test. In Wayland the class medians are practically equal to the standards in the Fourth and Fifth grades, well above in the Sixth grade, but very much below in the Seventh and Eighth grades. The chief change from the previous year is that the Wayland Sixth grade is now well above the standard instead of below it as formerly.

TABLE 17b. AVERAGE OPERATIONS
May, 1922

Grade	IV	V	VI	VII	VIII
Cochituate	41	63	50	61	76
Wayland	38	66	43	38	44
Standard	30	47	34	53	65

When the averages of the pupils' scores in the four fundamental operations are considered, the work of the Cochituate classes is well above the standard in every grade. This also applies to the Fourth, Fifth and Sixth grades in Wayland. In the Wayland Seventh and Eighth grades, however, the medians are considerably lower than the standards.

SUMMARY OF THE RESULTS IN THE ARITHMETIC TESTS

In all but two cases, the Cochituate medians in the Arithmetic Tests surpass the standards. Usually the difference in favor of Cochituate is a very large one, and many of the classes equal or exceed the standard average median of the next higher grade. This raises the question as to whether there is not too much emphasis on the teaching of Arithmetic in the Cochituate school. It is probable that the Cochituate children are, as the Intelligence Tests indicate, somewhat superior in native ability to children in many other localities, and so should be able to do better work. However, there is a point beyond which drill and practice in arithmetic is not economical. It is quite certain that no additional time or attention need be given to arithmetic in this school, and it might possibly be well to lighten the requirements.

The results at Wayland are not quite so consistent. In general, the medians of the Fourth, Fifth and Sixth grades practically equal or exceed the standards. In the Seventh and Eighth grades, however, the standards are considerably higher than the median performances of the Wayland classes, with the exception of the Seventh grade median in the Problem Test. In general it may be said that the Wayland medians are below those of the Cochituate classes, though there is an exception in the Subtraction Test.

It does not seem likely that the low results in the Seventh and Eighth grades at Wayland are to be explained on the ground of inferior native ability on the part of the pupils, as the Intelligence Tests show these pupils to be on practically the same mental level with the pupils in Cochituate. The difference may be due to differences in the efficiency of the teaching, or to differences in the courses of study.

As the tests and standards had been revised since the previous testing, it is possible to make only very general comparisons with the work in 1921. It may be said, however, that Cochituate maintains its previous high level of work, while conditions at Wayland have improved in that there are now three grades obtaining medians above the standards rather than only two as formerly.

A very interesting point is brought out by reference to the study of the results of the Intelligence Tests. In 1921 it was pointed out that the Wayland Sixth grade could not be expected to attain the standards in any subject because the Intelligence Tests showed the mental level of the pupils in that grade to be rather low. In the second testing, however, it was found that the new Sixth grade was a decidedly superior one, and as has been shown above, the Sixth grade medians were this time above the standards. This illustrates very well the desirability of securing some concrete evidence on the mental make-up of the pupils who take subject matter tests. It shows the injustice of evaluating the work of a teacher on the actual accomplishments of the class after she has taught them without taking into consideration the kind of material she had to work with.

2. The Reading Test

The scoring of this test is so arranged that a standard percentage distribution of scores is offered to which the performance of any grade from the Second to the Eighth inclusive may be compared. This standard distribution is shown in Table 18a, together with the percentage distribution of the scores made in each grade in the two schools. It is not to be expected that small classes like these will give smooth distributions, but if a class has average ability half the cases or 50 per cent. should obtain scores above 50. Another way of comparing these scores

with the standards is to examine the medians of the several classes. The method of standardization is such that the average class should have a median score of 50.

TABLE 18a. PICTURE SUPPLEMENT SILENT READING TEST

Form 2, May, 1922

Per Cent. of Pupils Receiving Each Mark

	Cochituate							Wayland					
Grades	III	IV	V	VI	VII	VIII	Standard	III	IV	V	VI	VII	VIII
100										8	18		
98							1—						
92						7	1+						
86	8				8		3						
80			3		8	7	4				9		7
74	4.			6		7	6				9	8	
68	8	6	6	19	13	26	8			25	27	8	7
62	4	13	9	19			10	25		17	9		13
56	4	19	9	6	15		11	17	9	8		8	20
50	20		3			33	12			8		17	
44	13	25	12	6		7	11	17	18				
38	25	13	15	13	15		10	17	46	8	9		13
32	4	19	12	13	8	7	8	8			9	8	20
26	8		20		23	7	6	17	9	8	9	17	13
20			6	19	8		4		9	17		17	7
14			3				3					8	
8		6	3				1+						
2							1—					8	
0									9				
No. Cases	24	16	34	16	13	15		12	11	12	11	12	15
Median Score	50	47	40	53	43	55	50	47	41	62	71	32	43

According to these methods of comparison all the grades in Cochituate except the Fifth and Seventh are up to the standard in this test. In Wayland the Third grade practically reaches the standard, and the Fifth and Sixth grades surpass the standards for these grades by considerable margins, but the other three grades fall much below. The score in the Seventh grade is an especially low one.

There has been a marked improvement in the reading since the previous testing, as shown by the following comparison of median scores.

	Cochituate						Wayland					
Grade	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
1921	38	56	20	25	26	32	26	32	38	20	20	38
1922	50	47	40	53	43	55	47	41	62	71	32	43

In 1921 the only grade median which equalled the standard was that of the Fourth grade at Cochituate, while the medians in the other eleven grades were extremely low. In 1922 only five of the medians were below the standard, and even in these five grades there has been a marked improvement. This change alone would seem to justify the first survey which pointed out the deficiency in reading ability.

3. The Penmanship Test

In Table 19a are shown the medians of the various grades in the Speed of Writing test, together with standards and the results obtained in three other Massachusetts communities. It will be seen from the table that the Cochituate medians are, except in the Fourth grade, somewhat higher than the Standards. They also compare very favorably with the results obtained in the three systems listed in the table for comparative purposes. The Wayland medians are slightly below the standards in grades II, V and VII, and somewhat above in grades VI and VIII. The Fourth grade median in this school is the only one which is materially below the standard. In comparison with the other three school systems, Wayland shows well in grades V, VI and VIII, but falls considerably below them in grades III, IV and VII.

TABLE 19a. SPEED OF PENMANSHIP
May, 1922

Grade	III	IV	V	VI	VII	VIII
Cochituate	48	54	80	99	94	96
Wayland	42	46	74	91	82	101
Standard	49	62	76	87	90	93
Newton	55	59	73	85	94	102
Brookline			76	87	90	
Fall River				86	89	

A comparison of the work in the two years is shown below.

	Cochituate						Wayland					
Grade	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
1921	60	61	77	85	100	113	48	58	68	63	88	98
1922	48	54	80	99	94	96	42	46	74	91	82	101

The losses are slightly more numerous than the gains, although there are no large losses except in the Third and Eighth grades at Cochituate. The Sixth grade in each school shows a rather large gain—14 letters per minute at Cochituate, and 28 letters per minute at Wayland. The latter gain is probably explained when we remember that the Wayland Sixth grade in 1921 was shown to be somewhat inferior mentally.

Table 20a shows the results of the Quality Test compared with the standards and the results from the three school systems whose medians were used for comparison in the Speed Test. No median in any grade of either school equals or exceeds the standards of the results of the other systems. In practically every case the difference is a large and significant one.

TABLE 20a. QUALITY OF PENMANSHIP
May, 1922

Grade	III	IV	V	VI	VII	VIII
Cochituate	28	32	33	38	37	39
Wayland	33	31	35	38	36	41
Standard	36	39	44	46	47	49
Newton	50	45	48	51	50	53
Brookline			44	46	47	49
Fall River				44	47	

Neither school has done as well in this test as in 1921, as may be seen from the comparison of the grade medians for the two years shown below.

	Cochituate						Wayland					
Grade	III	IV	V	VI	VII	VIII	III	IV	V	VI	VII	VIII
1921	39	46	43	45	50	53	32	40	60	48	50	57
1922	28	32	33	38	37	39	33	31	35	38	36	41

Many of the losses are large, amounting to more than a full step on the Ayres scale, by which they were graded.

SUMMARY AND SUGGESTIONS

As the previous report pointed out, the results of such a survey as this must be used, in the main, by the teachers and administrative officers who are on the ground and have full knowledge of the local condition. However, there are certain points brought out in the preceding pages which may be commented upon by the surveyors before closing this report.

In the first place, it will be noted that there are still many differences between the performances of the children of the two schools in the various tests. When the first report was written it seemed possible to explain these differences largely on the basis of differences in native intelligence, but the 1922 testing seems to show that these variations were largely removed by the reorganization of the classes. It appears that some other reason must be sought for the differences now existing.

Generally speaking, it may be said that the work of the Cochituate children is superior to that which is done at Wayland, although there are exceptions to this general rule. The Cochituate medians are also generally equal or superior to the standards in the various tests, except in the Quality of Penmanship. The arithmetic scores are so far above the standards as to indicate the possibility that some of the time spent on this subject could be used with greater profit in other fields.

The Wayland results are an improvement over the previous year in arithmetic and reading. The Sixth and Seventh grades, however, are still below the standards in most cases.

Both schools made marked gains in the Reading Test. In 1921 only one grade median equalled the standard, while in 1922 only five grades fail to attain the standards, and each of these had made some improvement over the previous performance.

The results of the Penmanship Tests were not as satisfactory as those of 1921, since there were losses in both Speed and Quality. The Quality scores are decidedly too low, and efforts should be made to improve the handwriting. This should not be done at the sacrifice of speed, as the Speed scores are not too high.

There are many exceptions to these general statements both among the grades and among the individual pupils. These cases must be singled out by superintendent and teacher for the special attention which they need.

APPENDIX

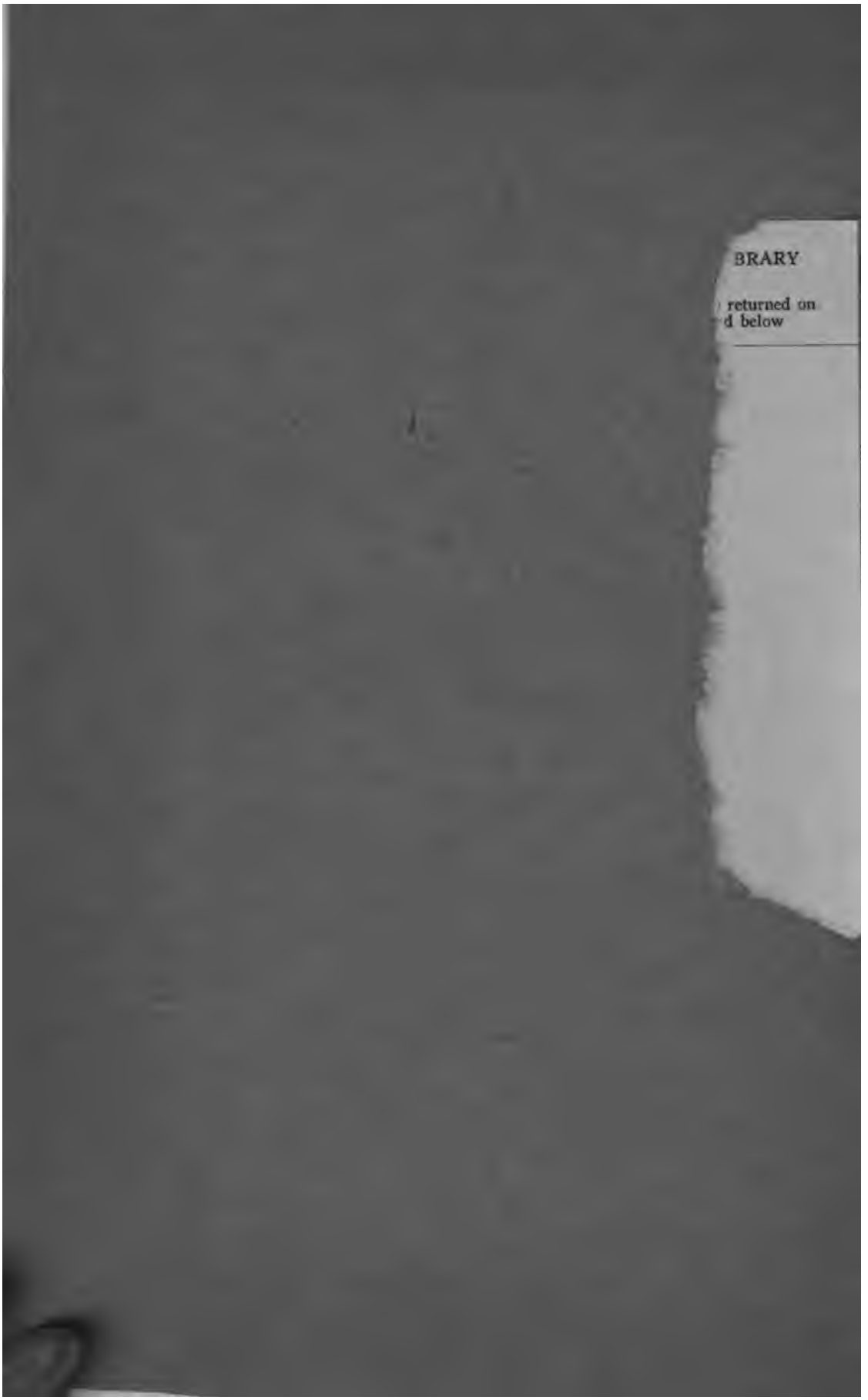
In order that the teachers might take full advantage of the results of the survey, tables like the following were appended showing the score of the individual pupils in each of the tests. A study of such a table by the teacher enables her to proceed intelligently in her efforts with individuals of her class.

WAYLAND and COCHITUATE

May—1922

COCHITUATE

Name	Age	Intelligence Examination	Mental Age	Arithmetic						Penmanship						
										Speed			Quality			
				Prob.	Add.	Div.	Sub.	Mult.	Ave. Oper.	1 min.	4 min.	Ave.	Speed	Dict.	Ave.	Reading
C., K.	10-5	49	15-2	84	39	61	103	86	72.3	87	93	90	21	26	24	60
D., R.	10-10	17	9-10	24	22	44	31	39	34	81	94	88	30	20	25	30
N., H.	10-1	31	12-2	72	89	101	71	85	86.5	86	85	86	45	29	38	54
P., R.	11-9	17	9-10	60	62	55	42	49	52	103	77	90	28	40	34	42
R., W.	9-11	18	10-0	40	63	79	61	49	63.	90	101	96	26	23	25	30
S., W.	11-10	33	12-6	56	74	39	52	34	49.8	87	87	87	23	26	25	24
W., H.	11-10	29	11-10	84	76	56	55	40	56.8	94	95	95	26	33	30	42
W., R.	14-2	33	12-6	48	46	77	88	91	75.5	88	85	87	30	36	33	42
Y., W.	12-8	20	10-4	26	63	39	50	60	53	106	80	93	20	26	23	48
S., C.	12-9	17	9-10	50	49	32	49	45	43.8	115	73	94	20	26	23	24
S., W.	10-6	32	12-4	24	78	87	50	78	73.3	114	92	103	20	25	23	48
E., E.	11-2	42	14-0	56	42	89	80	103	78.5	129	97	113	36	40	38	60
L., E.	11-6	29	11-10	44	65					No papers						
S., M.	10-8	29	11-10	74	104	105	77	101	96.8	77	78	78	38	30	34	30
S., A.	11-0	19	10-2	40	87	63	61	32	60.8	68	86	77	31	38	35	30
T., H.	11-6	53	15-10	78	89	130	114	103	109	112	70	91	35	60	48	72
W., L.	12-11	19	10-2	24	57	50	58	88	63.3	117	100	109	20	23	22	36
R., M.	11-5	25	11-2	58	130	69	89	52	85	66	78	72	33	36	35	36


BRARY
returned on
d below

Whole No. 2 Series 1, No. 2

HARVARD MONOGRAPHS IN EDUCATION

THE MARKING SYSTEM OF THE COLLEGE ENTRANCE EXAMINATION BOARD

BY
L. THOMAS HOPKINS
Graduate School of Education
Harvard University

Series 1 No. 2

STUDIES IN EDUCATIONAL PSYCHOLOGY
AND
EDUCATIONAL MEASUREMENT
Edited by
WALTER F. DEARBORN

OCTOBER, 1921

Published by
THE GRADUATE SCHOOL OF EDUCATION
HARVARD UNIVERSITY, CAMBRIDGE 38, MASS.

This study was undertaken at the suggestion of Professor Walter F. Dearborn of the Harvard Graduate School of Education. The writer is greatly indebted to him for assistance and counsel during the progress of the investigation.

The Marking System of the College Entrance Examination Board

This study represents an investigation into the distribution of the marks of the College Entrance Examination Board for the years 1902 to 1920 inclusive. It was made in order to discover if there were any grounds for the strong criticism of the college entrance examinations by New England educators, more especially secondary school principals and teachers. It is published at this time because the Board in its Twentieth Annual Report recognized the existence of sudden and violent fluctuations, from year to year, in the results of the examinations, in many subjects, and voted to employ expert assistance to aid in determining the specific causes.

SCOPE OF THE STUDY.

The subjects selected were English Readings, Elementary French, Elementary Algebra and Plane Geometry for the reason that they were offered by nearly all candidates, thus involving a relatively large number of cases. The arrangement of marks has been altered somewhat. A sample distribution as published by the board is as follows:

Solid Geometry	90-100	75-89	60-74	50-59	40-49	0-39
1916/1152*	1.8%	6.1%	18.2%	12.8%	14.1%	47%

Most of the larger colleges and universities admit on a mark of 60 or above while some of the smaller institutions will accept as low as 50. Assuming that the distribution ought to approximate the normal, for reasons which will be established later, and that anyone rated below 50 has failed to pass, the data in each case have been corrected from the above to read as follows:

Solid Geometry 1916/1152	90-100	75-89	60-74	50-59	0-49
	1.8%	6.1%	18.2%	12.8%	61.1%

The highest number of cases involved in any distribution was Elementary Algebra 1920/5249 and the lowest Elementary French 1902/509 with only 13 out of the 76 instances when the number fell below 1000.

FACTS BROUGHT TO LIGHT.

The following significant facts were discovered:

(*a*) Out of 76 distributions graphed every one is bimodal with the exceptions of:

English Readings 1902/800, 1906/1380, 1907/1661, 1908/1698, 1912/1731.

---

* In this and all similar cases the numerator of the fraction represents the year and the denominator the number of persons taking the examination.

In every instance the second mode in the distribution occurs in the assignment of the lowest marks and very often contains a greater percentage of cases than the one in the middle.

(*b*) Every distribution is skewed negatively or toward the lower end of the distribution of marks except:

Elementary Algebra 1906/1180, 1913/1916, 1918/3826.
Elementary French 1909/1196, 1916/2872.
English Readings 1903/996.

(*c*) The order in which the subjects approximate the normal distribution is as follows: English Readings, Elementary French, Elementary Algebra, Plane Geometry. In Figs. I and II are reproduced twenty selected graphs, five for each of the above subjects respectively.

EFFECT OF YEARLY INCREASE.

Various reasons suggested themselves as to why the results are so far from those expected. Bimodal distributions usually indicate a poor selection of cases. As the second mode in every instance is in the lower end or failure group, this might be caused by the influx of a large number of unprepared persons in the hope of slipping by. This explanation is discarded, however, for (*a*) the data show that this does not occur at intervals but appears regularly in all subjects, (*b*) the yearly increase in the number of candidates, with the exception of 1916, has been relatively constant as is shown in Table I.

RECOMMENDED CANDIDATES.

If all candidates of doubtful preparation could be eliminated a different result might be obtained. Consequently graphs were made for the years 1912-1916 inclusive for "only those candidates who were recommended for examinations on the ground of full and satisfactory preparation."*

It was found, however, that

(*a*) In Elementary Algebra and Plane Geometry, every distribution is bimodal, seven out of every ten are skewed negatively or toward the lowest grades, while the other three are skewed positively or toward the highest grades.

(*b*) Of the five in Elementary French, four are bimodal and three are skewed positively.

(*c*) In English Readings only one, 1916/2431, is bimodal, all the others tending roughly toward the normal.

---

* Further study of the group could not be made, as only these limited data are published by the Board.

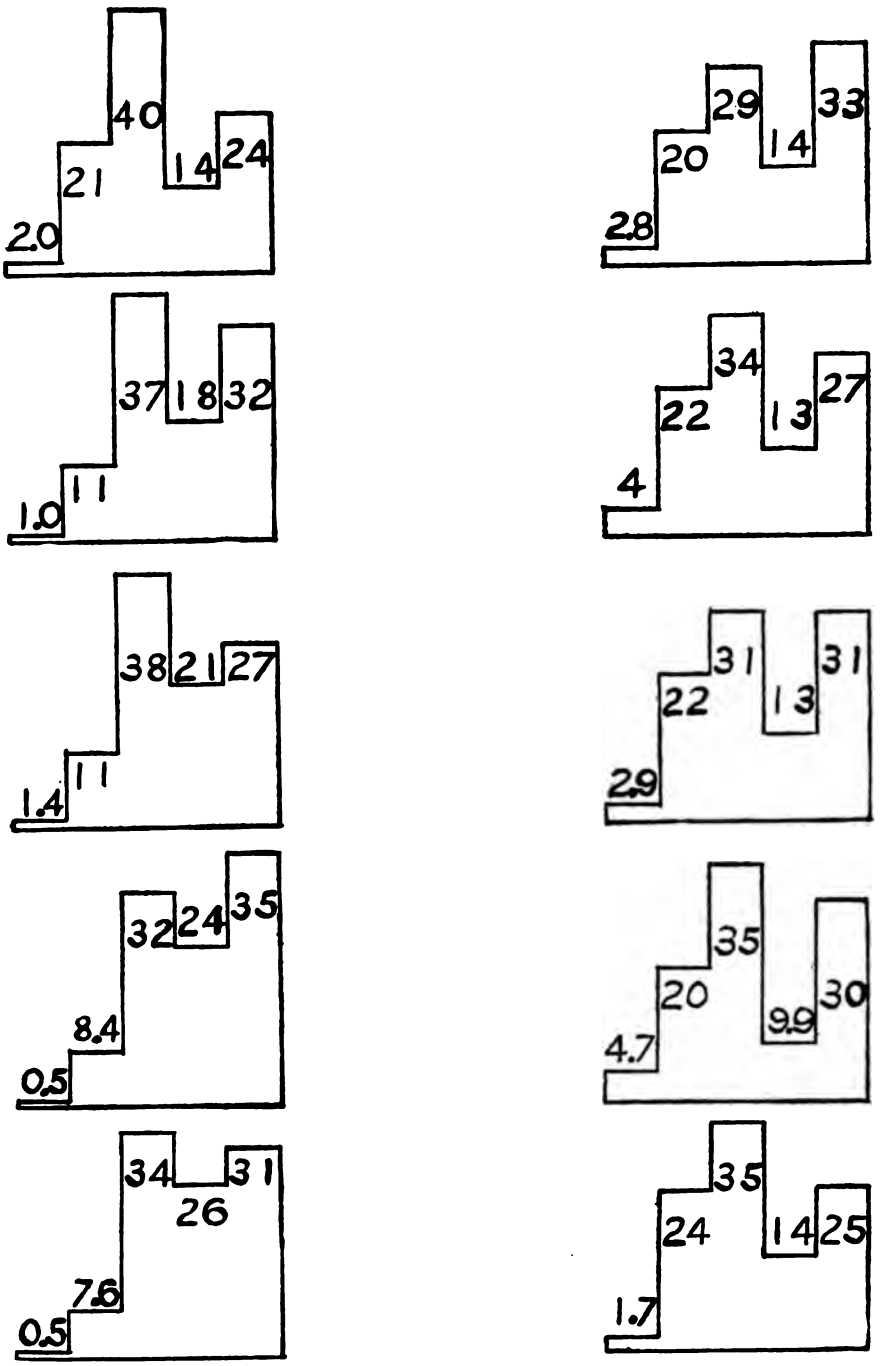


Fig. I—Graphs in the first column represent English Readings, the second Elementary French. The different divisions are as follows: 90-100, 75-89, 60-74, 50-59, 0-49. The figures show the percentage of cases.

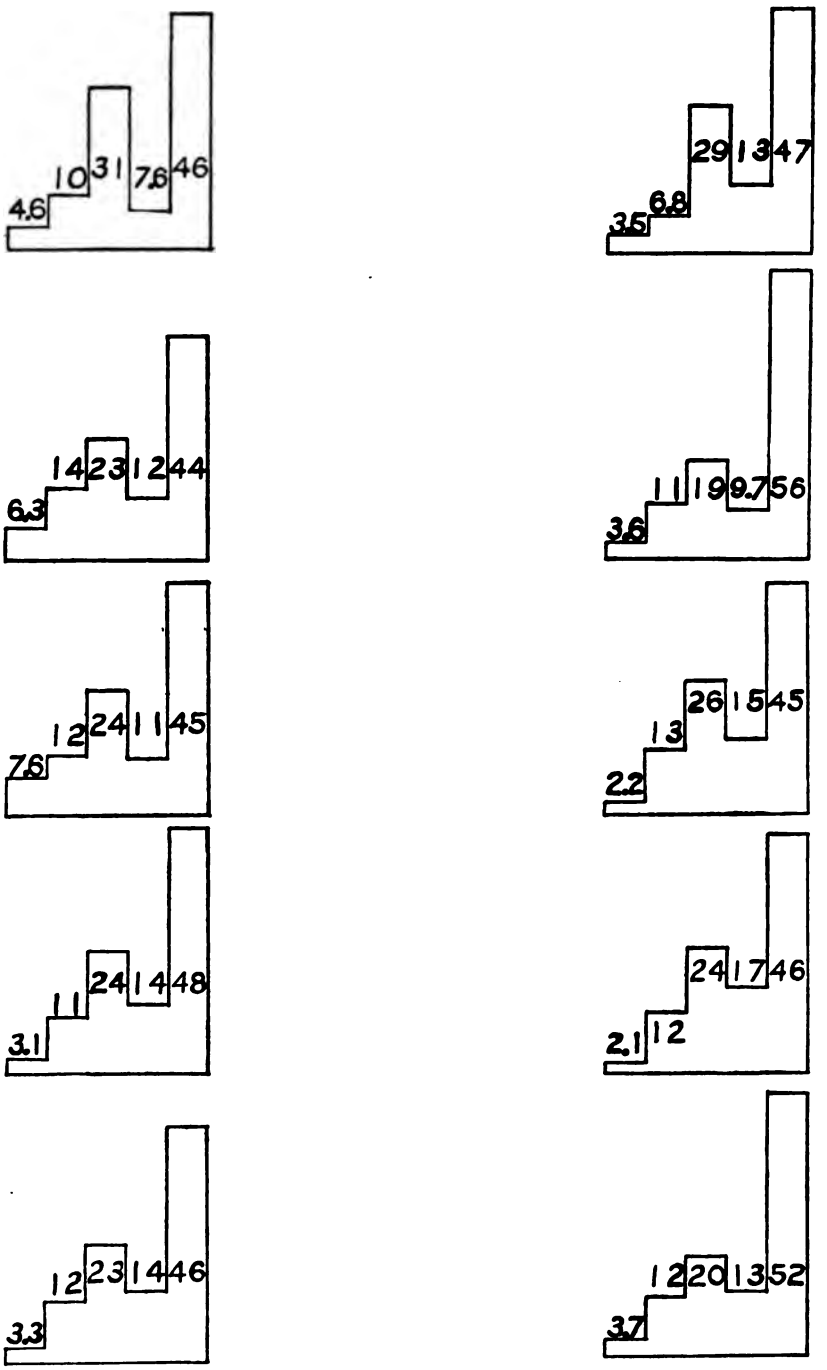


Fig. II—Graphs in the first column represent Elementary Algebra, the second Plane Geometry. Divisions as in Fig. I.

Table 1

Increase in the Number Taking Examinations

Year	English Reading		Elementary French		Elementary Algebra		Plane Geometry		In All Subjects	
	Total	Increase	Total	Increase	Total	Increase	Total	Increase	Total	Increase
1902	800		509		810		782		11744	
1903	996	196	625	116	973	163	927	145	14263	2519
1904	1033	37	661	36	1060	87	994	67	15275	1012
1905	1244	211	742	81	1079	17	940	54*	16189	914
1906	1380	136	854	112	1180	101	1069	129	17467	1278
1907	1661	281	1044	190	1291	111	1206	137	20034	2567
1908	1698	37	1143	99	1324	33	1171	35*	20607	573
1909	1706	8	1196	53	1445	121	1425	254	22208	1601
1910	1748	42	1166	30	1482	37	1340	85*	22189	19*
1911	1814	66	1317	151	1655	173	1586	246	22932	743
1912	1731	83*	1153	164*	1476	179*	1473	113*	20568	2364*
1913	1795	64	1299	299	1960	484	1743	270	22975	2407
1914	1963	168	1424	125	1233	727*	1833	90	23350	375
1915	1734	229*	1441	17	1380	147	1936	103	23990	640
1916†	4163	2429	2872	1431	3179	1799	3775	1739	47842	23852
1917	3327	836*	2284	588	2851	328*	3179	596*	37992	9850*
1918	3399	72	3211	927	3826	975	3832	653	41621	3629
1919	3582	183	3983	772	4181	355	4442	610	44406	2785
1920	2733	749*	4883	900	5249	1068	5227	785	48449	4043

NOTE:—*Decrease.

†The large increase in all subjects for 1916, as given in the Sixteenth Annual Report, is due chiefly to the joint action of Harvard, Yale and Princeton Universities, who agreed to discontinue their own June examinations and accept the results of the comprehensive papers prepared by the Board.

It is very evident from this that there is slight improvement in the ratings of the recommended candidates in English Readings and Elementary French but none in Elementary Algebra and Plane Geometry. The difference, however, is not marked enough to conclude that it is due to better preparation.

TOTAL YEARLY RANKS.

Theoretically, as the number of cases increases the nearer the distribution should correspond to the normal. Graphs were prepared showing the distribution of the total number of marks given for all subjects from 1902 to 1920 inclusive for all candidates, and from 1912 to 1916 for recommended candidates only. These show that in every case, (*a*) the distribution is bimodal, (*b*) it is skewed toward the lower end. Fig. III gives a selected list of graphical representations for totals of different years.

If all of the marks assigned in all subjects from 1902 to 1920 inclusive were combined into one grand total average distribution it would be as follows:

Grand Total	90-100	75-89	60-74	50-59	0-49
445,620	4.78%	18.34%	31.14%	13.78%	31.96%

In other words out of 445,620 cases only 4.78% received the highest grade while 31.96% failed. How many of the latter tried over again and succeeded there are no data to show.

A grand total average distribution for only those candidates recommended on the ground of full and satisfactory preparation as published for 1912 to 1916 inclusive is

Grand Total	90-100	75-89	60-74	50-59	0-49
87,642	6.35%	22.32%	32.28%	13.69%	25.36%

This is slightly better than the one given above, but considering the fact that the individuals involved here were highly selected, a failure of one-fourth, or 21,910 cases out of 87,642, places upon the Board the responsibility for a condition which is far reaching in its social and economic effects.

SELECTED DISTRIBUTIONS.

That the reader may have some samplings of extreme variations as a basis of comparison a selected list of graphs is given in Fig. IV. These are taken from different subjects and different years. The lowest number of cases involved is 641 while the highest is 2063.

WHAT WAS EXPECTED.

As was said at the beginning of this article, it was expected that the results would approximate the normal distribution. Briefly the evi-

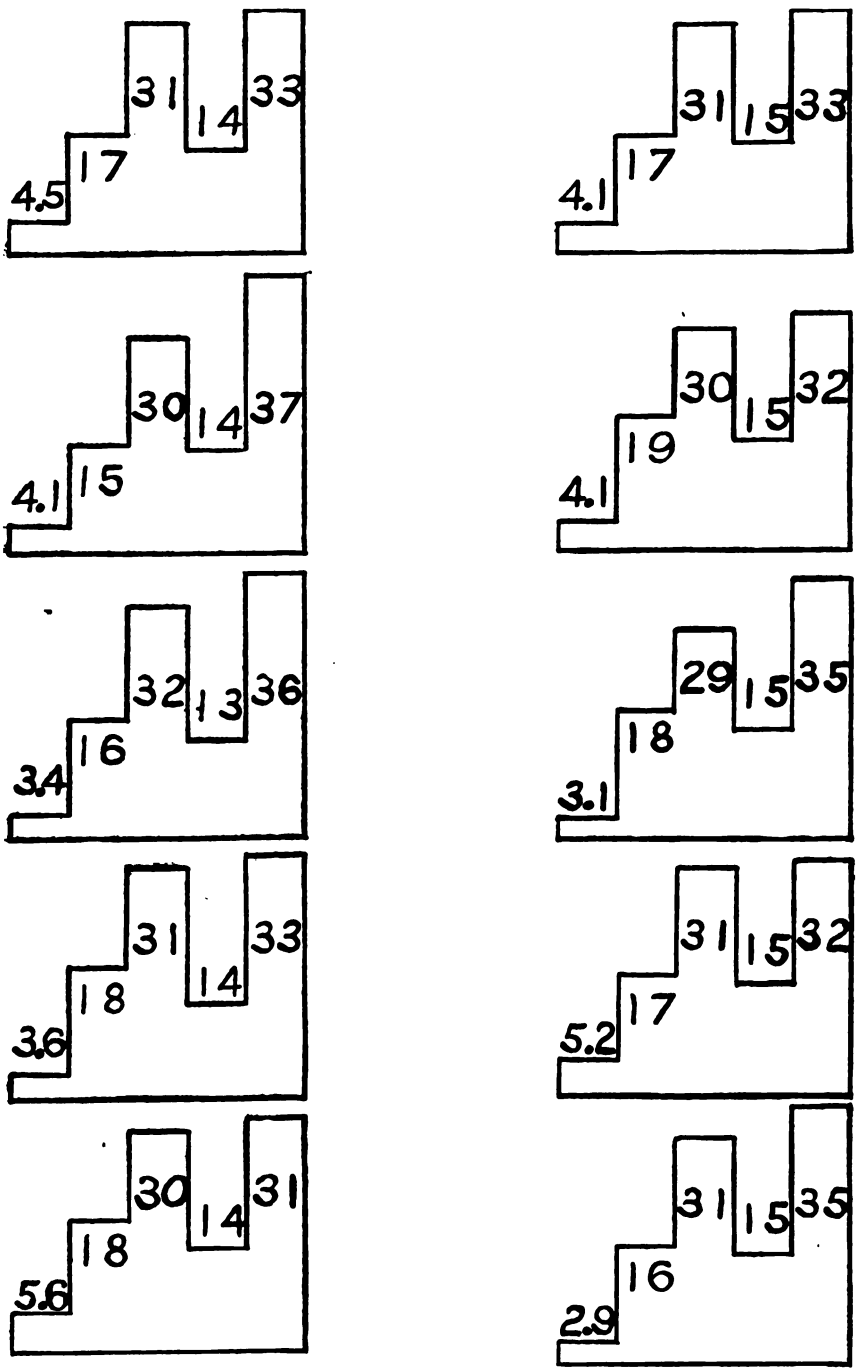


Fig. III—Totals for different years. Number of marks assigned will be found in Table I. Divisions as in Fig. I.

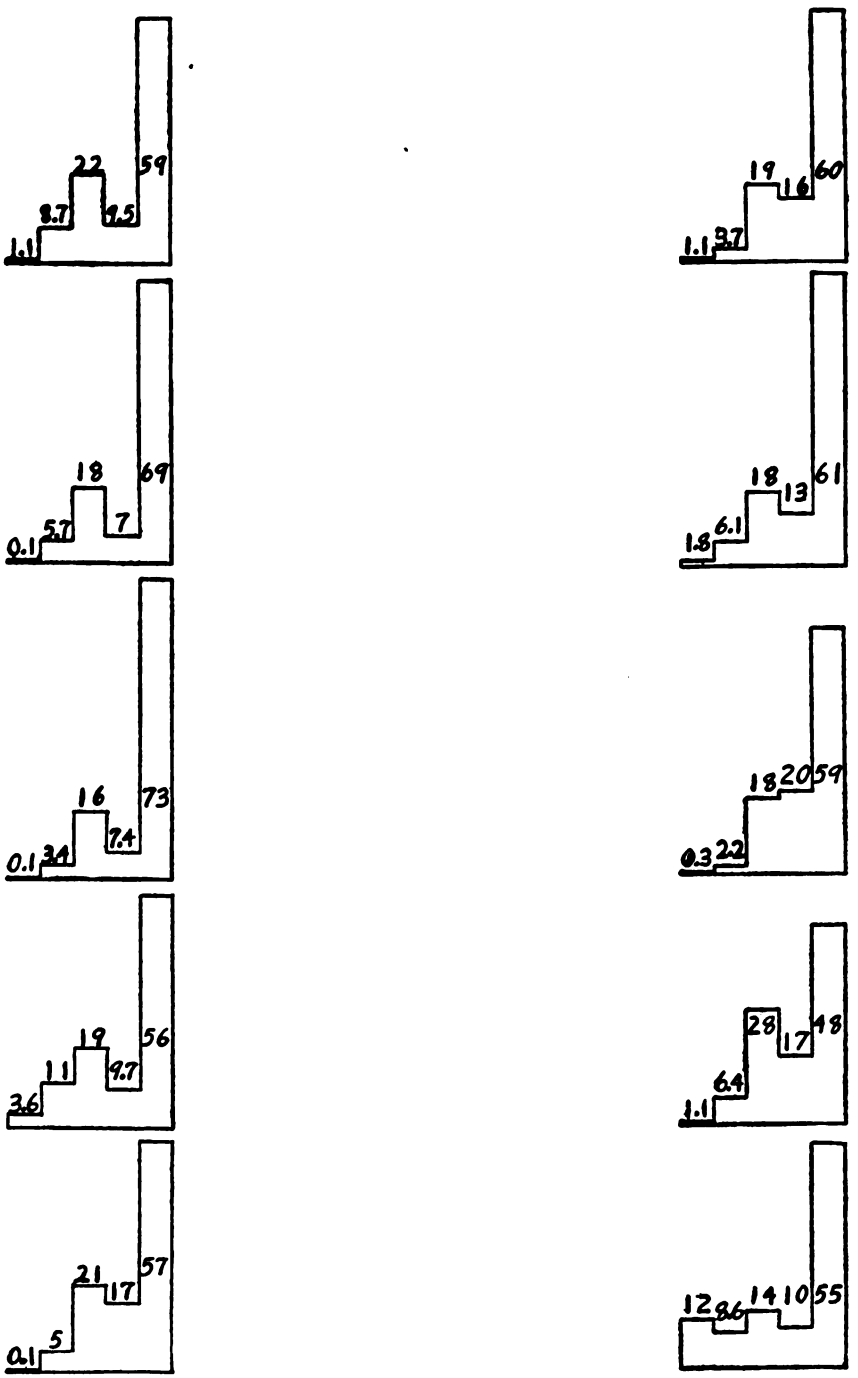


Fig. IV—Selected distribution in different subjects and years. The range of cases involved is from 641 to 2063. Divisions as in Fig. I.

dence supporting this is as follows: (a) Physical differences approximate the normal curve* as do mental characteristics,† (b) Marks, representing, as they do, estimates of mental abilities, are themselves distributed according to the same frequencies as the abilities they are designed to represent,‡ (c) The normal distribution of marks is the one usually found when a fairly large number of students are graded.§

Concluding then that the assignment of any relatively large number of grades ought to approximate the normal distribution and steadily so as the number increases over 500, this further question remains: What is the best method of dividing this distribution into groups for translating standing into a scale of marks? After a careful examination of all possible schemes we have concluded that the five division one is best. This is based on the orientation of a large number of cases around a central group whose accomplishment is considered median or average. Above and below lie groups of smaller size containing superior and inferior students in relation to the average and above and below these the still smaller groups of exceptions or failures.

The method of dividing our theoretical distributions into the five divisions which we will represent by the letters A, B, C, D, E, would be as follows: Find the median of the distribution and lay off on the base, on either side, the distance of 1 P. E. Within the area embraced by this $\pm$P. E. there will fall 50% of the total number of cases. This would represent the center or average or C group. Now lay off on either side of $\pm$ P. E. a distance equal to 2 P. E. Each one of the areas thus designated will contain 23% of these cases,‖ and would be represented by the letters B and D respectively. Again laying off the distance of 2 P. E. on either side we will reach the limits of the normal curve as for all practical purposes the ordinate may be taken as zero when the abscissa is 5 P. E. The last two divisions just made would each contain 2% of the total number of cases and would be represented by the letters A and E. The relationship between the cases represented by the five divisions of our normal probability integral and our marking system would now be as follows:¶

A	B	C	D	E
2%	23%	50%	23%	2%

* Brooks: The Foundation of Zoölogy, pp. 156-157, and Yule: An Introduction to the Theory of Statistics, p. 84.

† See the distribution of the IQ's of 905 unselected children 5-14 years of age in Terman: The Measurement of Intelligence, p. 66.

‡ Dearborn: School and University Grades. University of Wisconsin Bulletin No. 368.

§ Dearborn, Ibid, also Foster: The Administration of the College Curriculum, pp. 250-300.

‖ A table of the values of P. E. of the normal probability integral will be found in Rugg: Statistical Methods Applied to Education, p. 391.

¶ This was the division used by Buckingham in the standardization of the Buckingham Spelling Scale.

In like manner if we should lay off on either side of the mean the distance of A. D. we would find the following distribution:

A	B	C	D	E
2%	20%	56%	20%	2%

or if we should take for our unit .5σ and then lay off 1σ on either side our relationship would be as follows:*

A	B	C	D	E
7%	24%	38%	24%	7%

What is more commonly used by writers than either of the two preceding is to lay off the distance Q on each side of the mean. We would then have:†

A	B	C	D	E
3%	22%	50%	22%	3%

One of the first thoro treatments of variation in the marking of examinations was published by an English economist, Professor F. Y. Edgeworth, in the Journal of the Royal Statistical Society, September, 1888. This paper showed that there is a probable error of 3% and a possible error of 9%, in assigning a mark as representative of a student's real proficiency. Professor Edgeworth argued as a remedy that marks should be distributed according to the normal probability curve, but offered no suggestions as to its division. Many of the later writers, however, made definite divisions as given below:

		A	B	C	D	E
Cattell	(1905)‡	10%	20%	40%	20%	10%
Meyer	(1908)	3	22	50	22	3
Dearborn	(1910)	2	23	50	23	2
Foster	(1911)	3	22	50	22	3
Slosson	(1911)	3	22	50	22	3
Smith	(1911)	10	15	50	15	10
Ruediger	(1912)	4	24	44	24	4
Gray	(1913)	7	22	42	22	7
Cajori	(1914)	7	24	38	24	7
Starch	(1917)	7	24	38	24	7

* This was the division used by Ayres in the construction of the Ayres Spelling Scale.

† Tables of the values of AD, σ and Q of the normal probability integral will be found in Thorndike: Mental and Social Measurements, pp. 219, 220.

‡ Professor Cattell recognized the P. E. distribution of cases. He altered the percentages to more nearly meet the needs of classroom teachers who deal with small numbers, usually not exceeding 40.

A study of Figures 1 to IV inclusive will show no such relationship between the percentage of cases in the five divisions as is brought out here. Indeed one is amazed at the remarkable extent of divergence.

EFFECT OF READING METHODS ON THE DISTRIBUTION.

A number of examiners and readers have been consulted, from whom the following facts have been ascertained:

(*a*) Any paper marked between 50 and 60 by a reader is re-read by one or more before a permanent rating is given. This is due to the fact that the passing mark for some of the larger universities is 60 while that of many smaller colleges is 50. The re-examination of the paper is to determine whether the writer shows sufficient actual knowledge of subject matter and indicates enough potential possibilities of development to profit by the work offered in that department of a large university. If in the opinion of the examiner he does not, then the mark is below 60 which will admit only to the smaller colleges.

(*b*) Any paper marked over 90 by a reader is re-read by one or more readers before it is given its final mark. This is due to the fact that many prizes depend upon the highest awards.

(*c*) Any paper originally marked between 60 and 90 is never re-read except in rare instances when the rating is only a few points above 60.

(*d*) At the beginning the examiners agree on a value to be assigned to each question. There are two different methods of determining this. In some cases it is arrived at as follows: (1) Accepting 100 as the highest possible score, when there are ten questions each is given a value of 10. If there are eight questions each is given a value of 12½. When there are two or more parts to any question each part is given a proportion of the value assigned to the question as a whole, i. e. if there were ten questions the value of each would be 10. If one were divided into two parts, 5 would be given to each part. (2) In other instances the rating assigned is arrived at by taking the composite evaluation of each question by the readers. A clear exposition of this method as applied to French will be found in an article by Professor Donald C. Stuart of Princeton in the Bulletin of the New England Modern Language Association, September 1917.

That this method of reading the papers is a contributing cause of the poor distribution of marks is evident for, (a) no conferences are held between the examiners and readers to agree on the interpretation and value to be assigned to questions, (*b*) no attempt is made to standardize values of questions by considering the percentage of answers correct or incorrect, (c) the principle is not recognized that the assignment of

marks aggregating 1000 to 5000 in a subject, or 11,000 to 44,000 for a yearly total, ought to conform to the curve of error and hence no attempt is made to check up or correct results on the basis of the normal distribution.

CONCLUSION.

The facts seem to show clearly that, (a) only in rare instances, in the subjects studied, does the assignment of marks nearly approximate the normal, (b) the same condition holds true for the annual total for all subjects, (c) the results in cases where the pupils taking the examination are recommended by their school authorities on the ground of full and satisfactory preparation are only slightly improved, (d) this cannot be due to an influx of unprepared candidates as the increase in numbers each year is relatively constant and the poor distribution is found annually from 1902 to date, (e) the method of reading and scoring the papers, especially the lack of standardization of values and corrections in conformity with the curve of error, is a very natural factor in causing the existing conditions, (f) the suggestion is made that some approximation to the normal curve offers the best basis for solving present irregularities. This need not affect the passing marks as they may still be determined by such principles as govern them at the present time, altho a reconsideration of these might well be made by the Board.

Finally, in view of the large number of cases, no sufficient justification exists for the wide difference in the relative percentages assigned in the different subjects. Whether the distribution approximates the curve of error, or some other form, a certain uniformity in the different subjects may reasonably be expected. To accomplish this there must be co-operation between examiners and readers in the different subjects.

The writer wishes to emphasize the fact that this article does not claim to present an exhaustive study of the marks given by the College Entrance Examination Board. There are many phases of the subject which have not been touched. Sufficient evidence has been produced, however, to show the existence of an unwarranted condition and it is hoped the movement already inaugurated by the Board will result in a definite, workable plan for improvement.

SELECTED REFERENCES

Boring, E. G. — Marking System in Theory. Pedagogical Seminary 21: 269-77, 1914.

Breed, F. S. — Administering the Relative Marking System. School and Society, 5: 474-9, 1917.

Cajori, F. — A New Marking System and Means of Measuring Mathematical Abilities. Science, 39: 874-881, 1914.

Carter, R. E. — Correlation of Elementary Schools and High Schools. Elementary School Teacher, 12: 109-118.

Cattell, J. McK. — Examinations, Grades and Credits. Pop. Sci. Mon., 66: 367-78, 1905.

Dearborn, W. F. — Relative Standing of Pupils in the High School and in the University. University of Wis. Bull. No. 312.

——— — School and University Grades. University of Wis. Bull. No. 368

Dickson, J. D. H. — Percentages in School Marks. Nature 81: 367- 1909.

Edgeworth, F. Y. — The Element of Chance in Examinations. Journal of the Royal Statistical Society, 1890, 460-75, 644-73.

Ferry, Dean — Grading College Students. Williams College Bull., Series 8, No. 5, 1911.

Finkelstein, I. E. — The Marking System in Theory and Practice, Warwick and York, 1913.

Fisk, T. S. — Analysis of the Examination of 1911 of the College Entrance Examinations Board. Educational Review, 43: 155-67.

——— — Reports of the Secretary of the College Entrance Examination Board, 1902-1920.

Foster, W. T. — The Administration of the College Curriculum, Boston, 1911.

Holmes, Henry W. — The Teaching of Economics in Harvard University. Harvard Studies in Education, Vol. III.

Inglis, A. J. — Variability of Judgments in Equating Values in Grading. Educational Administration and Supervision, January, 1916.

Judd, C. H. — A Comparison of Grading Systems in High Schools and Colleges. School Review, 18: 460-70.

Lincoln, E. A. — Relative Standing of Pupils in High School, in Early College and in College Entrance Examinations. School and Society 5: 417-20, 1917.

Meyer, Max — The Grading of Students. Science, 28: 243-52.

Nickolson, F. W. — New Methods of Admission to College. Education, 32: 261-65.

Pettit, W. A. W. — Comparative Study of New York High School and Columbia College Grades. Teachers College, Master's Essay, 1912.

Roecker, W. F. — An Objective Study of the Rating of Traits in School Achievement. School Review, 22: 406-410, 1915.

Rugg, H. O. — Teachers' Marks and Marking Systems. Educational Administration and Supervision, February, 1915.

Russell, J. E. — Educational Value of Examination for Admission to College. School Review, 11: 42-54.

Sargent, E. B. — Education of Examiners, Nature 70: 63-68.

Sics, Raymond W. — Scientific Grading of College Students. University of Pittsburgh Bull., Vol. 8, No. 31, 1912.

Slosson, E. E. — A Study of Amherst Grades. Independent, 70: 836-39,

Smith, A. G. — A Rational College Marking System. Journal of Ed. Psy., 2: 383-93.

Smith, F. O. — A Rational Basis for Determining Fitness for College Entrance. University of Iowa Studies in Education.

Starch, D. — Can the Variability of Marks be Reduced? School and Society, 2: 242-43.

Thorndike, E. L. — Empirical Study of College Entrance Examinations. Science, 23: 839-45.

Young, W. H. — The High Schools of New England as Judged by the Certification Board. School Review, 5:15.

HARVARD STUDIES IN EDUCATION

Volume I: The Oberlehrer. A Study of the Social and Professional Evolution of the German Schoolmaster. By WILLIAM SETCHEL LEARNED. $1.25.

Volume II: The Appointment of Teachers in Cities. A Descriptive Critical and Constructive Study. By FRANK WASHINGTON BALLOU. $1.50.

Volume III: The Teaching of Economics in Harvard University. A Report Presented by the Division of Education at the Request of the Department of Economics. $2.00.

HARVARD BULLETINS IN EDUCATION

I. The School System as an Educational Laboratory. By WM. S. LEARNED. 25 cents.

II. Scales for the Measurement of English Composition. By FRANK W. BALLOU. 50 cents.

III. Bridging the Gap: The Transfer Class. By F. W. WRIGHT. 30 cents.

IV. A Selected Critical Bibliography of Vocational Guidance. By JOHN M. BREWER and ROY W. KELLY. 50 cents.

V. A Descriptive Bibliography of Measurement in Elementary Subjects. By HENRY W. HOLMES and others. 50 cents.

VI. Business Practice in Elementary Schools By ROY DAVIS. 50 cents.

VII. Sight-Saving Classes in the Public Schools. By R. B. IRWIN. 35 cents.

Published by

THE GRADUATE SCHOOL OF EDUCATION

HARVARD UNIVERSITY, CAMBRIDGE 38, MASS.

Whole No. 3 Series 1, No. 3
HARVARD MONOGRAPHS IN EDUCATION

STANDARD EDUCATIONAL TESTS IN THE ELEMENTARY TRAINING SCHOOLS OF MISSOURI

BY
WALTER F. DEARBORN
EDWARD A. LINCOLN
EDWIN A. SHAW

Series 1 No. 3

STUDIES IN EDUCATIONAL PSYCHOLOGY
AND
EDUCATIONAL MEASUREMENTS

Edited by
WALTER F. DEARBORN

JUNE, 1922

Published by
THE GRADUATE SCHOOL OF EDUCATION
HARVARD UNIVERSITY, CAMBRIDGE, MASS.

PREFACE

This investigation was undertaken for the Carnegie Foundation for the Advancement of Teaching to secure such evidence as school tests might furnish concerning the results of instruction in the Elementary Training Schools of the State Normal Schools and of the State University of Missouri; and to determine the relative attainments of the schools in these tests as compared with schools elsewhere, and also in comparison with each other.

Part of the material which appears in this study has already been published by the Carnegie Foundation in its report on "The Professional Preparation of Teachers for American Public Schools."* This report, however, was concerned chiefly with the results of the testing and did not emphasize the methods which were used in making the survey and working up the data. Several contributions to the technique of handling the test results were made during this investigation, which may be found useful in other school surveys; so that the republishing of the material, with the emphasis on methods rather than on results, seems desirable.

The report also has a certain historical interest as marking a transition from the observational method of estimating school achievement to the more objective methods of standard tests. At the time when this survey was undertaken,† standard tests had not yet come to be looked upon as a necessary part of such a project. The work of these schools was, therefore, at the same time carefully observed by several educational experts. How useful the tests would prove in supplementing, checking, or challenging the opinions of these observers was a matter about which neither the officers of the Carnegie Foundation nor the present writers were at all certain. The development, therefore, of what was then considered an ambitious plan of testing—eight tests besides the special reading tests were used—was frankly an experiment. The results, however, as the following statement from the official report indicates, furnished ample justification for the experiment. "These tests supplied an indispensable check upon the judgments of the observers, with which they tallied to a surprising degree."‡

*Carnegie Foundation for the Advancement of Teachers, New York, 1920.

†The first testing was done in May, 1916.

‡Professional Preparation of Teachers for American Public Schools, p. 5.

The writers are indebted to the officers of the Carnegie Foundation for permission to republish the parts of this study which were included in the above-mentioned report and for their cordial co-operation and assistance in the course of the survey.

W. F. D.
E. A. L.
E. A. S.

CONTENTS

§1. THE STANDARD TESTS: METHODS AND PROBLEMS OF MEASUREMENT

§2. INTER-SCHOOL COMPARISONS

STANDARD EDUCATIONAL TESTS IN THE ELEMENTARY TRAINING SCHOOLS OF MISSOURI

§1. THE STANDARD TESTS: METHODS AND PROBLEMS OF MEASUREMENT

The special problems which were met with in this survey, and which are discussed in the following pages, were in large measure due to the small numbers of children in the various groups compared. If the advantages which may result from the use of standard tests are not, however, to be limited to large school systems, these problems must be solved. That they were, for the most part, successfully met in the present instance will, it is believed, be borne out in the following report.

SELECTION OF TESTS

The tests employed were chiefly those in common use, in order that comparisons might readily be made with other schools and school systems. For the examination of reading, however, certain new tests were introduced, partly for the purpose of determining their availability for use in such surveys, and partly for such additional information as they might furnish.* Tests were given in reading, writing, arithmetic, spelling and composition.

GIVING THE TESTS

The tests were given in May, July, and October, 1916. Testing at these different times was necessitated in part by the fact that the sessions of the schools were not coincident. In order to note the possible effect of giving the tests at different times of the year, the school at Cape Girardeau, which was tested in May, was re-tested in October, and comparisons were made between the two sets of results obtained.

A COMPARISON OF TWO TESTINGS OF THE SAME GROUP

These comparisons are expressed in two ways. The central tendencies of the classes are compared, and then the results of correlation studies are shown. It must be held in mind that the former method of comparison is the most significant for the purposes of this report. The fact which this study wishes to determine is whether or not the *average* accomplishment of a class tested in October is approximately the same as the average accomplishment of the same class tested in the preceding

*The results of the special reading tests will be embodied in a later monograph.

May. The attainments of the individual pupils make little difference, for the comparisons of the schools and classes are made on the basis of the central tendencies of the groups compared.

The results are presented in the following tables.

In the Courtis Tests 21 out of 48 medians (44%) are exactly the same in the May and October results. Besides this, 17 October medians (37%) are only one example more or less than the medians of the same classes in May.

There is another way in which the central tendencies may be compared. This is by stating their differences in terms of their variations. For instance, if the May median is 9 examples done, with a P.E. of 1, and the October median is 10, it can be said that the latter is 1 P.E. greater than the former. This gives a more definite unit for comparison, and by the use of it all the comparisons are made on the same basis. In using this method for the purposes of this study the May medians were first taken as the starting points, and the relation of the October medians to them were expressed in terms of the P.E.'s of the former. Then the process was reversed, and the May medians were compared to those found in the October testing in terms of the P.E.'s of the latter. The two sets of results obtained in this way were then averaged.

TABLE 1. COMPARISON OF TWO TESTINGS AT CAPE GIRARDEAU*

a. Courtis Arithmetic Tests

1. *Addition*

Grade	3		4		5		6		7		8	
	"A"	"R"	"A"	"R"	"A"	"R"	"A"	"R"	"A"	"R"	"A"	"R"
May medians	3	0	4	0	6	3	8	3	8	5	9	7
P.E.	1	0	1	2	1	1	2	3	2	2	2	2
October medians	4	1	4	3	6	3	7	5	8	3	9	5
P.E.	1	1	1	1	1	2	1	2	2	2	1	2
2. *Subtraction*												
May medians	3	0	4	2	8	5	8	5	8	7	9	6
P.E.	1	0	1	1	1	1	2	3	2	2	2	3
October medians	3	0	4	2	6	5	6	4	8	6	10	7
P.E.	1	0	2	2	1	1	1	3	1	2	1	3
3. *Multiplication*												
May medians	1	0	3	0	5	3	8	4	7	4	10	6
P.E.	1	0	2	0	1	1	1	2	2	2	1	2
October medians	"O"	"O"	3	2	4	2	6	4	6	4	8	6
P.E.	o	o	1	2	1	1	2	1	2	2	2	4
4. *Division*												
May medians	"O"	"O"	1	0	4	1	5	3	5	4	8	7
P.E.	o	o	1	0	1	1	3	2	2	2	4	4
October medians	"O"	"O"	2	0	3	1	3	2	6	4	9	8
P.E.	o	o	1	0	1	1	3	2	2	2	3	3

*The eighth grade scores are not used in the comparison.

This second method shows that 64% of the medians in one test are within 1 P.E. of the corresponding medians of the other. Another 7% fall at plus 1 P.E., and 21% fall at minus 1 P.E.

It seems fair to conclude on the basis of these figures that the attainments of the classes were not very different in the two series of tests.

TABLE 2. COMPARISON OF TWO TESTINGS AT CAPE GIRADEAU

b. Stone Reasoning Test Scores

Grade	3	4	5	6	7	8
May medians	1.0	2.0	3.0	5.0	7.1	6.9
P.E.	1.0	1.0	1.0	2.0	1.0	3.1
October medians	1.0	2.0	5.0	5.6	7.8	9.2
P.E.	1.0	2.0	1.0	2.4	1.4	1.4

In the Stone test the medians are exactly the same in two grades, and in two others they are very close. This means that 66% of the medians fall within 1 P.E. of each other. The largest difference occurs in the eighth grade, but, as noted, no eighth grades are used in the comparisons.

TABLE 3. COMPARISON OF TWO TESTINGS AT CAPE GIRADEAU

c. Spelling Test Percentages

1. *Ayres List*

Grade	2	3	4	5	6	7	8
May medians	76	84	82	85	80	83	82
P.E.	14	8	13	12	10	13	12
October medians	73	93	89	92	85	83	92
P.E.	18	5	9	7	12	11	12

2. *Boston List*

May medians	57	65	66	57	55	62	64
P.E.	14	14	16	24	27	21	20
October medians	47	44	38	63	70	66	87
P.E.	27	11	18	12	15	14	11

The Spelling tests also give similar results. There are 54% of the medians which fall within the 1 P.E. limit, with 18% at plus 1 P.E. and 11% at minus 1 P.E.

TABLE 4. COMPARISON OF TWO TESTINGS AT CAPE GIRADEAU

d. Holmes Penmanship Speed Tests

1. *One Minute Period*

Grade	2	3	4	5	6	7	8
May medians	42	71	87	103	117	115	131
P.E.	5	5	9	8	7	12	15
October medians	44	60	69	84	100	111	107
P.E.	7	10	6	4	8	14	5

	2. *Four Minute Period*						
May medians	34	60	62	97	106	109	121
P.E.	6	8	10	8	8	9	9
October medians	40	58	72	79	103	113	98
P.E.	9	23	13	12	5	11	7
	3. *Average Speed*						
May medians	38	66	75	100	112	111	127
P.E.	5	3	7	5	7	9	15
October medians	43	59	70	85	102	114	104
P.E.	10	16	8	9	6	14	8

The figures for the speed of writing are not quite so satisfactory, but the medians are not far apart. Nearly half of them (45%) fall within 1 P.E., 18% are at plus 1 P.E., and another 11% are at minus 1 P.E. The first noticeable discrepancy occurs in this set of tests, for 14% of the medians are more than minus 2 P.E. apart, and the same number are separated by a distance of more than plus 2 P.E.

TABLE 5. COMPARISON OF TWO TESTINGS AT CAPE GIRADEAU

e. Holmes Penmanship Quality Tests

	1. *Quality on Speed Test*						
Grade	2	3	4	5	6	7	8
May medians	30	33	35	30	30	40	38
P.E.	3	5	3	3	3	5	9
October medians	30	38	35	45	38	33	45
P.E.	3	2	5	10	2	3	3
	2. *Quality on Dictation Test*						
May medians	35	38	47	53	48	53	60
P.E.	7	5	7	8	2	7	7
October medians	28	35	40	43	43	43	50
P.E.	3	3	5	7	7	5	2

The results in the tests for the quality of writing seem to be the most unsatisfactory of all. Yet even in these tests 32% of the medians are less than 1 P.E. apart, while 21% are at plus and 7% at minus 1 P.E.

TABLE 6. COMPARISON OF TWO TESTINGS AT CAPE GIRADEAU

f. Kansas Silent Reading Test

Grade	3	4	5	6	7	8
May medians	18.1	16.3	23.8	10.6	30.0	23.7
P.E.	9.5	4.6	4.8	2.4	4.2	7.5
October medians	20.7	17.7	20.7	22.4	35.9	34.6
P.E.	8.1	8.3	4.5	5.3	6.2	5.8

In this test the scores are rather close again. 42% of the medians are separated by less than 1 P.E., 25% are at plus 1 P.E., and 8% at minus 1 P.E. Only in the sixth grade is there a considerable difference.

TABLE 7. COMPARISON OF TWO TESTINGS AT CAPE GIRADEAU

g. Holmes Speed of Reading Test

Grade	2	3	4	5	6	7	8
May medians	147	384	237	222	234	463	444
P.E.	60	180	30	42	57	100	81
October medians	137	225	272	231	310	452	499
P.E.	94	38	62	90	52	69	19

Although the absolute differences are rather large in the speed of reading test, comparisons on the basis of the variations show that these differences are not so great as they seem. No less than 64% of the medians fall within the 1 P.E. limit, with 15% at plus 1 P.E., and 14% at minus 1 P.E.

TABLE 8. SUMMARY OF COMPARISONS OF TWO TESTINGS AT CAPE GIRADEAU

P.E.	—5	—4	—3	—2	—1	0	+1	+2	+3	+4	+5
Courtis Tests	0	0	1%	5%	17%	64%	7%	1%	0	0	0
Stone Test	0	0	0	17%	0	66%	9%	8%	0	0	0
Spelling Tests	4%	0	0	7%	11%	54%	18%	7%	0	0	0
Speed of Writing	0	0	2%	19%	3%	43%	12%	10%	7%	0	5%
Quality of Writing	3%	3%	3%	14%	25%	21%	18%	7%	0	0	4%
Kansas Silent Reading	17%	0	0	8%	8%	42%	25%	8%	0	0	0
Speed of Silent Reading		7%	0	0	7%	64%	15%	0	0	7%	0
Combined Tests	3%	1%	1%	10%	14%	52%	12%	5%	1%	1%	1%

Taken as a whole, the results seem to show that the differences between the medians obtained in the two testings were not very great. It seems fair to infer that the attainments of the classes as a whole were not much different in October from what they were in the preceding May.

With one exception, namely, in the quality of writing, the coefficients of correlation in Table 9 are sufficiently high to support the assumption that the work of the classes in the two tests had not changed appreciably. It is to be remembered that, in any repetition of tests, there are a good many factors other than changes in the relative attainments of the pupils which may affect the relative standing of the pupils, and that for the purposes of this report it is chiefly important that the average or median standings of the classes as a whole should not change materially.

TABLE 9. COMPARISON OF TWO TESTINGS AT CAPE GIRADEAU

Correlations of the First and Second Tests

	Test	n	r		P.E.
1.	Courtis Addition Rights	63	73.3	±	3.9
2.	Courtis Multiplication Rights	62	79.1	±	3.2
3.	Stone Reasoning Test	63	83.2	±	2.6
4.	Average Speed of Writing	77	89.6	±	1.7
5.	Quality of Handwriting on Speed Test	79	27.8	±	6.9
6.	Kansas Silent Reading	64	65.3	±	4.9
7.	Speed of Silent Reading	77	59.3	±	5.0
8.	Ayres Spelling List	76	69.3	±	4.0

From the above considerations, it seems fair to assume that neither the lapse of a few months' time nor the possibilities of practice gains had any appreciable effect on the results. In the tabulation of results all pupils were, therefore, listed as members of the grades in which they were enrolled in May.

PERSONNEL OF EXAMINERS

The testing was done by five graduate students* of the Division of Education of Harvard University. Special care was given to the training of the examiners in order that the methods of conducting the tests might be uniform. In the dictation of words for spelling, and in one of the reading tests, the examiners were assisted by selected teachers in the several schools.

TECHNIQUE OF TESTING

The sequence of the tests was arranged to guard in as far as possible against the influence of fatigue. Usually an interval of at least a class period intervened between the various tests. The Courtis and Stone tests were usually given on different days, and whenever possible in the morning hours. The schedule for testing was similar in each of the schools. The attitude of the pupils toward the testing was, it is believed, not essentially different from their attitude toward the regular school work.

*Messrs. H. B. Cummings, L. P. Damon, E. A. Lincoln, C. A. Puckett, and E. A. Shaw.

CORRECTION OF PAPERS

The correcting and scoring of all the papers were done by two, and in some instances, by three of the examiners. In those cases where the personal factor might enter into the rating, every paper was scored independently by each of the readers. The grading of the penmanship papers was done by three of the examiners who practiced until they had attained a high degree of uniformity in their judgments. After practice, their final ratings fell within one step in terms of the Ayres' scale. A similar method was followed with the compositions. All of the papers were read by two of the examiners who had through practice acquired an equal proficiency in rating the papers. When there were small differences in their judgments, the average was taken.

ADVANTAGES OF EXPERT READERS

It is believed that many errors which inevitably appear when the work of correcting is done by teachers and pupils have been eliminated by having a limited number of trained readers for the work. The skill and insight attained by the examiners through practice is accumulative. The time consumed in correcting is decreased, while the accuracy of the work is increased, and the results are interpreted with greater discernment and uniformity.

LIMITED SCOPE OF THE WORK

All that these tests assume to do is to determine the present attainments of the pupils now in the schools. It is quite possible that in some of the classes the pupils are drawn from an inferior school population. It was stated in one school, for example, that many of the pupils were recruited from a "floating" population. Their previous training may thus have been so limited that it would be unfair to hold the school in question responsible for their present attainments.

Some of the classes were very small. In such classes it is, of course, possible to have by mere chance relatively inferior or relatively superior groups. The facts in the case might have been determined, e.g., by supplementary mental tests, had such studies lain within the scope of this investigation.* Some indication of the character of these classes has, however, been secured from a study of the age-grade distribution. In some cases it is evident that these small classes are not representative

*Had the group methods for giving intelligence tests been developed at the time when this survey took place, it would have been an easy matter to determine the important facts about the fundamental abilities of the pupils in the several schools and classes. The extent to which differences in the mental abilities of pupils and classes of pupils may affect the standard test scores has been clearly shown since group intelligence examinations became available. See, for example, School and Society, Vol. XII, No. 306, pp. 441-444.

of the grade of work which is being done; but certain conclusions can be safely made about them, as will be shown later.

Since there are no ninth grades at Kirksville and Maryville, results applicable to the eighth grade could not be obtained in the October testing. The results, as tabulated, are for that part of the eighth grade which was present in May.

PRESENTATION OF RESULTS

The results are stated in the tables in terms of the averages or medians, depending generally on the method of presenting standards with which it is desired to make comparison in this study. It is usually stated that the median cannot be used when the groups are very small, but that the averages must be chosen as the central tendency most representative of such groups. As the finding of the median is much easier and more rapid than finding the average, the data were examined to see if the use of the former could be justified.

The results of this investigation are given, in part, in the accompanying table, No. 10. In this table the averages and average deviations are compared with the medians and median deviations of the same grades in the same tests. The selection of the schools and grades for this comparison was a random one except for the fact that the smaller grades were chosen, as it is in the small groups that averages are supposed to be the most accurate measure.

The table clearly shows that the use of the median was amply justified. In many cases the two scores are exactly the same, and in many more the difference is so slight as to be negligible. In no instance does one of the scores fall outside the probable error or average deviation of the other. The only discrepancies which attract attention are in the speed of silent reading. But here an examination of the facts shows that the median is likely to state the tendency of the group better than the average. The reason for this is that a very high or a very low score has a decided effect on the average in the small group, but does not affect the median. A study of the distributions shows that this is what has happened in the cases where there is a marked disagreement between the average and the median.

TABLE 10. COMPARISON OF AVERAGES AND MEDIANS

	Penmanship		*Silent Reading*			
	Av. Speed	Av. Quality	Speed	Reproduction	Questions	
Average	39	33	189	21	41	
A.D.	8	5	94	14	13	22
Median	38	33	147	14	39	Pupils
P.E.	5	5	60	9	14	Grade 2

Average	33	44	169	33	51	
A.D.	7	7	65	9	9	10
Median	32	42	147	37	53	Pupils
P.E.	12	6	33	9	4	Grade 3
Average	57	39	184	34	40	
A.D.	13	6	60	18	20	16
Median	59	39	180	40	49	Pupils
P.E.	14	4	60	17	8	Grade 4
Average	81	45	321	50	70	
A.D.	15	6	84	12	8	8
Median	83	45	363	57	72	Pupils
P.E.	16	9	81	12	7	Grade 6
Average	82	50	285	51	67	
A.D.	17	8	80	9	12	20
Median	79	48	299	51	72	Pupils
P.E.	18	8	67	11	13	Grade 5
Average	92	41	310	56	69	
A.D.	15	9	78	8	15	12
Median	89	37	291	57	69	Pupils
P.E.	18	7	72	9	14	Grade 7

The variability from the central tendency, i.e., average or median deviation, is stated under the average or median. This figure, taken together with the number of scores from which the central tendency is calculated will indicate the reliability of the averages or medians. Besides the central tendencies of the separate schools, grade averages and medians with their variabilities have been calculated for the combined grades of all the schools taken together. These figures give a basis for further comparisons.

METHODS OF INTER-SCHOOL COMPARISON

Several methods of comparing the relative standing of the various schools and grades have been employed. This has been necessary partly because of the small number of cases involved in most of the comparisons, but also because of other considerations. The average or median standing of the various grades has usually been used in the inter-grade comparisons; but, in the inter-school comparisons since the work of the upper grades is more representative of what the schools have accomplished than that of the lower grades, the standing of the higher grades may properly be given greater weight. Since our tests of the eighth grades are sometimes incomplete, the average or median standing of

the sixth and seventh grades taken together has been secured for this comparison. This method of comparison has been called the *Upper Grade Median Method.*

The average or median of the two grades has been taken rather than that of either grade alone, because the increased number of individuals in the combined grades gave figures that were more reliable, and also because this practice guards to some extent against the effect of one or the other of the grades being by chance exceptionally good or poor. When one of the grades was much larger than the other a random selection of cases was made from it, in order that the combined grade medians might be based upon approximately equal numbers from each grade.

The above method of ranking did not always prove adequate to express all the facts, and in questionable cases several other methods have been used. The supplementary methods employed are the following:

(1) *Progress Method:* Comparison of standing of pupils in each grade with that of the pupils in the preceding grade, with special reference to the standing of the seventh grade as compared with that of the second or third grade. Obviously, the efficiency of a school can be judged in part by the evidence of superior or inferior accomplishment on the part of the latter grades as compared with the earlier grades of the school. It is true that the attainments, e.g., of an exceptional third grade when compared with an inferior eighth might seem to indicate that the school had not accomplished much, and this possibility limits somewhat the usability of the method. The average age of the pupils in such a third grade in comparison with the age of the pupils in the eighth grade may in some instances explain this result. The Progress Method has, therefore, always been supplemented by a study of the age-grade relationships. If, for example, the average age of the pupils in the various grades increases uniformly, but the relative attainments of the pupils advances more rapidly, this can usually be taken as evidence of the superior efficiency of training rather than as evidence of the mere maturity of the pupils. Some very striking evidence of the close relationship between age and school attainment will be pointed out later. For the present it is enough to say that the relationship seems so close that, when there is some irregularity in school progress, a reference to the age-grade table will almost invariably explain the difference. Obversely, the striking instances in this report of the relative superiority of a succeeding grade over a preceding one can in most cases be explained by a reference to the age grade table.

2. *Rank Sum Method:* By this method each of the grades is considered separately. For example, the second grades of all the schools studied are ranked according to their relative standing in a given test. Next, the third and the subsequent grades are handled in the same way. Then the relative standing of the school as a whole is judged by the sum of the first, second, third, etc., ranks secured by the various grades.

This method has the advantage of giving weight to the attainments of the earlier grades, and thus it supplements the first two methods which rate the schools chiefly on the basis of the attainments of the latter grades. It should be said, however, that it is a fair assumption that relative superiority in the earlier grades is likely to be indicative of the selection of a superior group rather than excellence of school training; but this is not always true. Therefore, although it is believed that the first two methods outlined above are on the whole better, it has seemed fair to give some weight in judging the school as a whole to the accomplishments of the pupils of the earlier grades.

(3) *Distribution Method:* At times the best results are obtained by a direct study of the distributions or frequency tables of the results of the tests. These give a record of each individual score, and thus often throw considerable light on the above-mentioned problems of comparison.

The following tables will show two rankings. The first of these will be based on the scores of the combined sixth and seventh grades; the second and final ranking will take into consideration all the other methods of comparison above described.

In order that the reader may fully understand the use of these methods of comparison, the following example is presented, which shows how the final ranking was obtained in the Stone Arithmetic test:

TABLE 11. DISTRIBUTION OF SCORES IN STONE ARITHMETIC TEST

School	0	1	2	3		4		5		6		7		8		9	10	11+		No. Cases			Med.
																				VI	VII	Tot.	
Cape Giradeau	0	1.4	2.0	3.0				5.0	5.6		6.6	7.0	7.6	8.0	8.8	9.0							
			2.0	3.0				5.0			6.6	7.2		8.0		9.2				13	13	26	6.6
				3.0				5.2			6.6	7.2											(1.4)
								5.4			6.6												
											6.8												
Columbia			2.0	3.0		4.0	4.6	5.0	5.6	6.0	6.6	7.0	7.6					11.2		11	14	25	5.0
				3.0			4.6	5.0		6.4													(1.6)
				3.0				5.4		6.4													
				3.0				5.4		6.4													
				3.0																			
				3.0																			
				3.0																			
				3.4																			
Kirksville		1.0	2.0	3.0	3.6	4.0		5.0	5.6	6.4	6.6		7.8	8.0	8.8		10.2	11.2		22	18	40	5.0
		1.0	2.0	3.0		4.0		5.0	5.6		6.6												(1.6)
		1.0	2.0			4.0		5.0	5.6														
		1.0	2.0			4.0		5.0	5.6														
		1.0	2.0			4.0		5.2	5.6														
			2.0			4.0		5.4	5.8														
Maryville		1.0	2.0	3.0		4.0		5.4		6.4	6.6	7.8						11.0	13.2	8	8	16	5.4
						4.0		5.4		6.4	6.6												(1.2)
						4.0																	
						4.0																	
Springfield	0		2.4	3.0		4.0	4.6	5.0	5.6	6.4	6.6	7.0	7.6	8.0	8.6	9.0	10.2			15	17	32	5.8
	0			3.0		4.0	4.6	5.2	5.8		6.8	7.0	7.6			9.2							(1.8)
						4.2	4.6	5.4					7.6			9.2							
							4.6						7.8										
Warrensburg	0	1.0	2.0	3.0	3.6	4.0		5.0	5.6	6.4	6.6		7.6	8.0	8.8		10.2	11.2	11.8	21	21	42	4.4
	0	1.0		3.0		4.0		5.4			6.6				8.8		10.2		12.2				(2.2)
		1.0		3.0		4.0		5.4			6.6												
		1.0		3.0		4.0					6.6												
		1.0		3.0		4.0					6.6												
		1.0				4.0					6.8												
						4.4																	

TABLE 12. GRADE MEDIANS IN STONE TEST

Grade	III		IV		V		VI		VII		Rank	Gain of VII
School	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Sum	over III
Cape Giradeau	1.0	**2**	2.0	**2.5**	3.0	**3**	5.0	**1**	7.1	**1**	9.5	6.1
Columbia	1.0	**2**	2.5	**1**	4.6	**1**	4.6	**2.5**	5.2	**5.5**	12.0	4.0
Kirksville	0	**5**	1.0	**4.5**	1.0	**6**	4.0	**4.5**	5.2	**5.5**	25.5	5.2
Maryville	0	**5**	0.5	**6**	3.1	**2**	4.0	**4.5**	6.7	**3**	20.5	6.7
Springfield	1.0	**2**	2.0	**2.5**	2.8	**4**	4.6	**2.5**	6.8	**2**	13.0	5.8
Warrensburg	0	**5**	1.0	**4.5**	2.0	**5**	3.6	**6**	5.6	**4**	24.5	5.6

From the foregoing tables the schools will be seen to rank as follows by the several methods.

Upper Grade Median	Progress	Rank Sum
Cape Giradeau	Maryville	Cape Giradeau
Springfield	Cape Giradeau	Columbia
Maryville	Springfield	Springfield
Columbia	Warrensburg	Maryville
Kirksville	Kirksville	Warrensburg
Warrensburg	Columbia	Kirksville

Cape Giradeau stands first by two methods of comparison, and is second only by a small margin in the third, so is placed first in the final ranking.

Springfield and Maryville are not far apart, but there is considerable evidence for placing the former school in second place. Springfield stands ahead of Maryville by both the Upper Grade Median and Rank Sum methods, and a comparison of the distributions shows a skew toward the upper scores by the Springfield pupils.

Columbia is given the fourth place over Kirksville because of the excellent Rank Sum score, and the superior performance in the lower grades.

An examination of the distributions seems to indicate that Warrensburg should probably not be ranked below Kirksville. The Progress and Rank Sum methods both favor the former, though only slightly. Probably these two schools should rank equally.

NUMBER OF PUPILS TESTED

It has been noted that the small number of pupils tested limits the scope and accuracy of the comparisons and conclusions of the report. The number of pupils is shown in Table 13 by schools and grades. There are several grades which contain less than 10 pupils, and very few of them contain more than 20. It must be remembered also that these figures express the maximum number of pupils tested. Because of absences, the number of papers obtained in any given test was often one or two less than the number appearing in the table.

Columbia has no eighth grade, and, as there are no ninth grades at Maryville or Kirksville, no results for the eighth grade could be obtained in the October testing of this school. For these reasons the results from the eighth grades have not been used in the inter-school comparisons.

Besides the eighth grades mentioned above the second and sixth grades at Maryville, and the third grade at Cape Giradeau were so small that the results are very unreliable.

TABLE 13. NUMBER OF PUPILS BY SCHOOLS AND GRADES

Grade	2	3	4	5	6	7	8	Total
Cape Giradeau	22	9	13	15	13	23	16	111
Columbia	12	10	19	11	13	14	0	79
Kirksville	19	16	17	13	25	19	9	118
Maryville	6	10	10	16	8	12	10	72
Springfield	16	16	16	12	16	17	14	107
Warrensburg	15	21	15	20	21	41	32	165
Grade Totals	90	82	90	87	96	126	81	652

AGES OF PUPILS TESTED

A study of the ages is important because of the light it throws on the question of the mental calibre of the classes. Other things being equal, it is likely that a grade composed of older pupils is an inferior grade, and conversely, a grade composed of younger pupils is likely to be a superior one. For valid comparisons, then, the median ages of the pupils in the grades, schools, or systems compared should be the same.

It is, however, plain from the table of ages (Table 14) that this condition is not fulfilled. The medians of the combined grades progress regularly, and with equal deviations, beginning at 8 years in the second grade, and reaching 14 in the eighth. But in the separate schools this regularity is found only in Columbia. In the seventh grade at Cape Giradeau the median is a year below that of the other seventh grades. That is, half the pupils are 12 years old and under, while in each of the other schools half of the seventh grade pupils are 13 years of age or under. Kirksville has one over-age grade, the fifth, in which more than half the pupils are a year older than the combined grade median. In Maryville, the fifth, seventh, and eighth grades are all above the median age.

Springfield presents the most interesting state of affairs. From the second to the sixth grades, inclusive, the pupils are a year under age as compared with the combined grade medians. But the seventh grade pupils are two years older than those of the sixth; and the eighth grade pupils are two years older than those of the seventh. It is also noticeable that most of the variations in the Springfield medians are

zero, which means that half the class are exactly at the median age. These facts seem to indicate that the upper grades in this school are somewhat inferior when compared with the rest of the school. Assuming this to be so, Springfield will suffer in comparison with the other schools, as it is on the basis of the upper grade scores that the comparisons are chiefly made. An examination of the results seems to show that this is exactly what happens. In many of the tests the Springfield seventh grade will be found to have scores decidedly lower than those of the sixth grade. In Warrensburg there are two grades, the fourth and the sixth, in which the median ages are a year below the combined grade medians.

The age problem may be handled in another way by studying the retardation and acceleration of individuals in each grade. The following tables show the findings on this basis. A pupil was considered retarded if he was two or more years older than the group median age for his grade. He was considered accelerated if he was two or more years younger than the grade median.

TABLE 14. MEDIAN AGES OF PUPILS BY GRADES*

Grade	2	3	4	5	6	7	8
Cape Giradeau	8	9	10	11	12	12	14
	(0)	(0)	(1)	(1)	(1)	(1)	(1)
	22	*9*	*13*	*15*	*13*	*23*	*16*
Columbia	8	9	10	11	12	13	
	(0)	(1)	(0)	(1)	(1)	(1)	
	12	*10*	*19*	*11*	*13*	*14*	*0*
Kirksville	8	9	10	12	12	13	14
	(1)	(1)	(1)	(1)	(0)	(0)	(1)
	19	*16*	*17*	*13*	*25*	*19*	*9*
Maryville	8	9	10	12	12	14	15
	(1)	(1)	(1)	(1)	(1)	(1)	(1)
	6	*10*	*10*	*16*	*8*	*12*	*10*
Springfield	7	8	9	10	11	13	15
	(0)	(0)	(1)	(0)	(0)	(0)	(1)
	16	*16*	*15*	*12*	*16*	*14*	*14*
Warrensburg	8	9	9	11	11	13	14
	(0)	(1)	(1)	(1)	(1)	(1)	(1)
	15	*21*	*15*	*20*	*21*	*41*	*31*
Grade Medians	8	9	10	11	12	13	14
	(1)	(1)	(1)	(1)	(1)	(1)	(1)
	90	*82*	*89*	*87*	*96*	*123*	*80*

*The figures in parenthesis express the median deviations, and the figures in italics tell the number of cases in each grade.

TABLE 15. NUMBER OF PUPILS TWO OR MORE YEARS RETARDED

Grade	2	3	4	5	6	7	8	All Grades
Cape Giradeau	0	0	0	3	3	1	3	10
Columbia	0	0	1	0	3	0		4
Kirksville	4	1	4	3	6	1	2	21
Maryville	2	1	3	3	1	3	3	16
Springfield	0	0	0	0	0	0	1	1
Warrensburg	0	2	1	5	5	6	4	23

TABLE 16. PER CENT OF PUPILS TWO OR MORE YEARS RETARDED

Grade	2	3	4	5	6	7	8	All Grades
Cape Giradeau	0	0	0	20	23	4	19	9
Columbia	0	0	5	0	23	0	—	5
Kirksville	21	6	24	23	24	5	22	19
Maryville	33	16	30	19	6	25	30	21
Springfield	0	0	0	0	0	0	7	1
Warrensburg	0	10	7	25	24	14	13	14

In Kirksville, Maryville, and Warrensburg there is a fairly large amount of retardation throughout the grades. The first two schools have retarded pupils in every grade, the latter in every grade but one. In Cape Giradeau the retardation comes in the last four grades. In both Springfield and Columbia the retardation is negligible except in one grade of the latter school. In Columbia this grade is the sixth. This fact will probably affect the standing of the school in the "upper grade median" method of inter-school comparison.

TABLE 17. NUMBER OF PUPILS TWO OR MORE YEARS ACCELERATED

Grade	2	3	4	5	6	7	8	All Grades
Cape Giradeau	0	0	1	1	1	4	2	9
Columbia	0	0	0	0	0	1	—	1
Kirksville	0	0	1	1	2	0	0	4
Maryville	0	0	0	1	0	2	0	3
Springfield	1	1	1	1	0	1	0	5
Warrensburg	0	2	1	0	9	5	2	19

TABLE 18. PER CENT OF PUPILS TWO OR MORE YEARS ACCELERATED

Grade	2	3	4	5	6	7	8	All Grades
Cape Giradeau	0	0	8	7	8	17	12	8
Columbia	0	0	0	0	0	7	—	1
Kirksville	0	0	6	8	8	0	0	4
Maryville	0	0	0	6	0	16	0	4
Springfield	6	6	6	9	0	7	0	5
Warrensburg	0	10	7	0	43	12	7	12

It is apparent that except in the case of Warrensburg, acceleration in these schools is not nearly as frequent as retardation. Such as occurs is concentrated for the most part in one or two grades. This fact may possibly affect the inter-school comparisons in Maryville and in Cape Giradeau, where the acceleration is fairly large in the seventh grades, and in Warrensburg, where there is acceleration in the sixth grade.

To conclude: The study of the ages of pupils points out probable differences in the mental make-up of the pupils in the different schools, and of pupils of the different grades in the same school, which must be kept in mind when the results from the various schools are compared.

§2. INTER-SCHOOL COMPARISONS

1. Arithmetic Tests

Two series of tests were given in Arithmetic; first, the Courtis Tests, Series B, and second, the Stone Reasoning Tests.

a. *The Courtis Tests.* The Courtis tests, Series B, consist of a set of twenty-four examples for each of the operations in arithmetic: Addition, Subtraction, Multiplication and Division. There is a definite time limit for each set. When the test is given each child is provided with a folder containing the four sets of examples, and is carefully instructed as to what he is to do. At a signal from the tester, all the children begin work on the first set of examples, and they continue until time is called. The same method of procedure is used for each set in turn. Between the second and third sets a short rest period is given the children.

The results of these tests in the Missouri Training Schools are presented in the accepted form for the purpose of comparing them with those of other schools. The results are shown in the body of the report in terms of "Speed" and "Accuracy" in accordance with the more recent tables of Courtis. The scores in the Attempts and Rights, as used in the method hitherto employed, are given in the tables in the appendix. The difference in these two methods of stating results is that the Rights are now calculated in terms of percentages of the number of attempts, and these percentages are termed "Accuracy." "Speed" and "Attempts" may be used interchangeably to indicate the number of examples tried in the allotted time.

The findings in the four fundamental processes of Addition, Subtraction, Multiplication, and Division are stated separately, and a rank in all four of the processes combined has been attempted.

Speed of Addition. The schools will first be ranked in terms of "Speed," or, what is the same thing, in the number of examples attempted. Tables giving the complete results of this and the following tests will be found in the appendix. The tables in the body of the text will embody the main findings only.

TABLE 19. SPEED OF ADDITION

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	3	4	6	8	8	8	1	1
Columbia	2	3	5	7	6	6	5	4
Kirksville	4	4	5	6	6	6	5	5.5
Maryville	5	6	6	7	8	7	2.5	2
Springfield	2	4	6	6	6	6	5	5.5
Warrensburg	3	5	6	6	8	7	2.5	3
All Schools	3	4	6	6	7	6		

As may be seen from the table, Cape Giradeau leads with 8 examples attempted. Maryville and Warrensburg are tied with 7 attempts each. The variability of the Warrensburg median is, however, twice as great as that of Maryville, and since the latter school ranks decidedly superior to Warrensburg by the Rank-Sum method of comparison, it seems that Maryville should have the second place.

In the table, Columbia, Kirksville, and Springfield all rank the same, each having 6 attempts as a median in the combined sixth and seventh grades. From an examination of the distributions it appears that Columbia should be ranked slightly ahead of the other two.

The results in Springfield are striking in that there is no progress after the fifth grade. This suggests the possibility of an inferior seventh grade—a possibility which has already been indicated by the study of the age-grade distributions. This school also has a smaller variability than Kirksville, but on the other hand, the Rank-Sum method shows that they are not materially different.

With all the above considerations in mind it seems that the ranking by the upper grade medians should be modified slightly, and that the final ranking should be as follows: (1) Cape Giradeau; (2) Maryville; (3) Warrensburg; (4) Columbia; (5.5) Springfield and Kirksville.

Accuracy of Addition. In the accuracy of addition Cape Giradeau easily stands in first place, while Springfield and Warrensburg come second and third in the order named. Columbia is placed fourth by the Progress method and because of the good score in the seventh grade. The supplementary methods show the Maryville record better than that of Kirksville. The absence of regular progress in this test, except in the results at Cape Giradeau, is noteworthy and quite the opposite of what might be reasonably expected.

TABLE 20. ACCURACY OF ADDITION

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	0%	20%	43%	60%	67%	63%	1	1
Columbia	0	0	50	33	60	41	5	4
Kirksville	0	25	20	38	50	40	5	6
Maryville	25	67	50	60	40	40	5	5
Springfield	0	25	50	50	64	56	2	2
Warrensburg	0	25	38	33	50	44	3	3
All Schools	0	20	40	44	50	50		

By consulting the corresponding table in the Appendix (Table II), it will be noted that the variability of these median scores is so large in the lower grades as to make them unreliable. This limitation does not

apply so seriously to the scores of the upper grades, although it is still a factor in some schools. It should be kept in mind when examining all the "accuracy" tables.

TABLE 21. SPEED OF SUBTRACTION

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	3	4	8	8	8	8	2.5	2.5
Columbia	1	1	4	6	7	7	4.5	5
Kirksville	1	4	6	5	6	5	6	6
Maryville	2	5	6	9	9	9	1	1
Springfield	4	4	7	8	8	8	2.5	2.5
Warrensburg	3	4	7	5	9	7	4.5	4
All Schools	2	4	6	6	8	7		

Speed of Subtraction. In this test all the methods give nearly the same ranking to the schools. Maryville stands first, with Cape Giradeau and Springfield tied for second place. Warrensburg is placed fourth and Columbia fifth because the latter stands lower by the Rank-Sum method. Kirksville is decidedly behind the other schools.

It is interesting to note the results in Columbia which show that the low standing in the lower grades where the formal arithmetic is not emphasized does not hinder rapid and regular progress afterward with a fairly high score in the last grade.

TABLE 22. ACCURACY OF SUBTRACTION

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	0%	50%	63%	55%	70%	63%	4.5	4
Columbia	0	0	50	43	86	83	1	1
Kirksville	0	0	33	40	60	40	6	6
Maryville	0	80	60	55	67	63	4.5	5
Springfield	25	20	71	75	67	70	3	3
Warrensburg	0	40	60	67	78	77	2	2
All Schools	0	20	60	57	75	67		

Accuracy of Subtraction. Columbia has a substantial lead in this test because of a remarkably high median score in the seventh grade. This is an example of how a single superior grade may affect the results. Cape Giradeau is placed above Maryville because it leads by both the Rank-Sum and Progress methods of comparison. Kirksville stands far behind the other schools.

Again we find little regularity of progress in the various schools, with the exception of Kirksville and Warrensburg. The scores in Maryville decrease from the fourth grade through the sixth.

TABLE 23. SPEED OF MULTIPLICATION

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	"0"	3	5	8	7	8	1.5	2
Columbia	"0"	1	1	3	4	4	5.5	6
Kirksville	"0"	1	3	3	6	4	5.5	5
Maryville	3	5	5	7	8	8	1.5	1
Springfield	"0"	3	5	8	6	6	3.5	3.5
Warrensburg	"0"	3	6	5	7	6	3.5	3.5
All Schools	"0"	2	5	5	6	6		

Speed of Multiplication. Maryville does better than Cape Giradeau by the Progress and Rank-Sum methods, and so is given first place in the final ranking. Springfield and Warrensburg stand very close together. Possibly Warrensburg is slightly in the lead, but not enough to place it a rank ahead. The supplementary methods indicate that Kirksville should be ranked ahead of Columbia. The zero scores in the lower grades mean simply that multiplication is not taught in these grades.

TABLE 24. ACCURACY OF MULTIPLICATION

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	0%	0%	67%	63%	75%	67%	1	1
Columbia	0	0	0	0	57	33	5	5
Kirksville	0	0	25	0	33	20	6	6
Maryville	0	25	75	50	50	50	3.5	3.5
Springfield	0	0	80	67	57	60	2	2
Warrensburg	0	0	50	50	60	50	3.5	3.5
All Schools	0	0	50	50	50	50		

Accuracy of Multiplication. There seems to be no good reason for changing the initial ranking. Maryville leads Warrensburg by the Rank-Sum method of comparison, but the latter school has a decidedly better score in the seventh grade. The large number of zero medians in the third and fourth grades mean that little work in multiplication is done in these grades. This holds true of all the grades up to the seventh in both Columbia and Kirksville. Absence of regular progress

is more noticeable in this test than in those preceding. There is no school in which regular and constant improvement is made throughout the grades.

TABLE 25. SPEED OF DIVISION

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	"0"	1	4	5	5	5	3.5	4
Columbia	"0"	"0"	1	2	5	3	5.5	5
Kirksville	"0"	1	2	3	4	3	5.5	6
Maryville	"0"	3	4	5	6	6	1.5	1
Springfield	"0"	2	3	6	5	6	1.5	2
Warrensburg	"0"	2	5	4	6	5	3.5	3
All Schools	"0"	1	4	4	5	4		

Speed of Division. By the Progress and Rank-Sum methods of comparison Maryville leads Springfield, and is therefore ranked in first place. The supplementary methods also place Warrensburg over Cape Giradeau in third place. A study of the distributions shows that the results in Columbia are better than those in Kirksville. The sudden rise in the seventh grade median in Columbia is notable. It probably means that the previous low scores are due to lack of stress on the work. The progress in Maryville and Kirksville is interesting in that each grade does, on the average, one more example than the preceding grade.

TABLE 26. ACCURACY OF DIVISION

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	0%	0%	33%	75%	80%	86%	1	1
Columbia	0	0	0	50	60	50	4.5	5
Kirksville	0	0	0	33	50	33	6	6
Maryville	0	60	57	50	60	50	4.5	4
Springfield	0	0	67	71	75	71	3	3
Warrensburg	0	0	60	67	80	80	2	2
All Schools	0	0	50	57	71	67		

Accuracy of Division. In the final ranking only one change has been made. Maryville is placed over Columbia because of a better standing by the Rank-Sum method of comparison. As was the case in Multiplication, the zero medians indicate that the schools do not teach Division in the lower grades to any extent. Progress is more uniform

in this test, as a regular gain is found in every school except Maryville. In the case of the latter school there is a decreasing score in the three middle grades.

In order to get some idea as to the relative standing in the four operations as a group, a table has been made which shows the rank attained by each school in each test. A ranking in the combined operations has also been attempted.

TABLE 27. COMPARISON OF RANKS IN COURTIS TESTS

School	Addition		Subtraction		Multiplication		Division		Four Operations combined	
	S*	A*	S	A	S	A	S	A	S	A
C. Giradeau	1	1	2.5	4	2	1	4	1	2	1
Columbia	4	4	5	1	6	5	5	.5	5	5
Kirksville	5.5	6	6	6	5	6	6	6	6	6
Maryville	2	5	1	5	1	3.5	1	4	1	4
Springfield	5.5	2	2.5	3	3.5	2	2	3	3	2
Warrensburg	3	3	4	3	3.5	3.5	3	2	4	3

It will be seen from the table that the correlation between the standings in the different tests is not a perfect one. There is, however, a decided tendency for a school to hold approximately the same rank throughout the tests in both speed and accuracy, and for this reason the final ranking is probably valid.

A further ranking may be made by combining the standings in speed and accuracy. In this Cape Giradeau stands first. Springfield and Maryville are practically tied for the second place, but the higher rank should probably go to the former school on the basis of greater accuracy in the work. Warrensburg, Columbia, and Kirksville follow in the order named.

b. *Stone Reasoning Test.* In this test the pupils are given a folder which contains twelve problems of graded difficulty. Exactly fifteen minutes are allowed the pupils in which to work as many problems as they can. A certain problem value, as determined by the author of the tests, is credited for each correct answer. The pupil's score in the test is obtained by adding all the credits which he receives.

The following table gives the results in terms of the median scores.

*S—Speed. A—Accuracy.

TABLE 28. MEDIAN SCORES IN STONE REASONING TEST

Grade	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	1.0	2.0	3.0	5.0	7.1	6.6	1	1
Columbia	1.0	2.5	4.6	4.6	5.0	5.0	4.5	4
Kirksville	0.0	1.0	1.0	4.0	5.2	5.0	4.5	5.5
Maryville	0.0	0.5	3.1	4.0	6.7	5.4	3	3
Springfield	1.0	2.0	3.6	4.6	6.8	5.8	2	2
Warrensburg	0.0	1.0	2.0	3.6	5.6	4.4	6	5.5
All Schools	1.0	1.6	3.0	4.0	6.4	5.2		

The differences between the schools in this test are not great, and in most cases do not amount to more than the median variations of the scores. In some of the lower grades the scores are not at all reliable because of the great variations.

Cape Giradeau has a slight advantage over Springfield for first place. The latter is very nearly tied with Maryville. The distributions show Springfield somewhat superior, and this school also leads by the Rank-Sum method, while the Progress method favors Maryville. The two schools are evidently not far apart, but the Springfield score is a little more reliable because it is based on a larger number of cases.

From a study of the distributions it also appears that Warrensburg, Columbia and Kirksville differ but little. Columbia is given the fourth place because of the excellent showing made by the Rank-Sum method. The Warrensburg score is not very accurate, as is indicated by the large median deviation and a bimodal distribution.

There is much better progress from grade to grade in nearly every school than was shown in the Courtis tests, and in this the combined grade medians show especially well.

2. Spelling Test

For these tests thirty words were selected for each grade. Twenty of these were taken from the Ayres' Spelling Scale, and the other ten from the Boston Minimum Lists.

The words were not given by the examiners, but were dictated by the regular grade or practice teacher. This was deemed advisable because of the fact that the pupils were more familiar with the voice and enunciation of the teacher.

In the correction of the papers credit was given for accepted forms of simplified spelling, and for wrong words, evidently misunderstood, which were correctly spelled. The necessity for the latter practice could be obviated by using sentences containing test words as a means of giving the examination.

It should be noted in this test that progress through the grades is not made by getting a higher score on the same words, but by maintaining the same score on lists of words of increasing difficulty.

The comparisons are made in terms of averages instead of medians because the standards are so stated.

TABLE 29. PER CENT. OF CORRECT SPELLING OF WORDS FROM AYRES' LIST

Grade	2	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	76	84	82	85	80	83	80	1.5	2
Columbia	not given	60	57	66	67	76	72	5	5
Kirksville	67	66	53	73	64	58	61	6	6
Maryville	97	93	73	80	76	76	75	3	3
Springfield	95	85	72	89	81	79	80	1.5	1
Warrensburg	82	65	58	76	79	71	74	4	4
All Schools	81	74	64	78	74	73	76		

a. *Words from Ayres' List.* The standing of the schools is shown in the above table. Springfield ranks slightly higher than Cape Giradeau in the distributions and by Rank-Sum method. The latter school stands ahead by the Progress method, but this is only because the earlier grades in Springfield make such good records.

The differences between Columbia, Maryville, and Warrensburg are very small; by the Upper Grade Median method these schools are practically tied. By the other methods of comparison employed, however, the third place falls to Maryville by a small margin. Warrensburg seems to have a slight advantage over Columbia, although this may be questioned. The Kirksville results are more than a grade behind those of any of the other schools.

TABLE 30. PER CENT OF CORRECT SPELLING OF WORDS FROM THE BOSTON LIST

Grade	2	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	57	65	66	57	55	62	63	1	2
Columbia	not given	38	28	35	53	36	44	4	5
Kirksville	36	33	29	59	39	33	33	6	6
Maryville	73	59	38	57	59	37	42	5	4
Springfield	73	61	71	71	74	50	62	2	1
Warrensburg	59	47	31	65	60	43	52	3	3
All Schools	57	50	42	58	55	45	49		

b. *Words from Boston List.* The relative standings of the schools in the spelling of words from the Boston lists is shown in the above table. The reason for the lower scores on the Boston List is that the words are much more difficult than those of the Ayres' list, and the Boston standards were obtained from tests given after the words had been studied by the pupils.

Here again the Upper Grade Median method of comparison is not alone sufficient, and the final ranking must be made by the other methods. These give the first place to Springfield. The fourth place is given to Maryville instead of Columbia because the former school is found to lead decidedly by the Rank-Sum method, and by a study of the distributions. The distributions show that the mode of the combined sixth and seventh grades at Maryville is 10% higher than that of the combined sixth and seventh grades at Columbia. Kirksville is again a grade or more behind the other schools. It is worth noting that the records of some of the schools are better in the lower grades than in the upper. This is especially noticeable in Kirksville, Maryville and Warrensburg.

The average deviations are nearly all very high; some, indeed, so large as practically to invalidate the averages. This means that the range of ability in spelling the words on the Boston list was very great, or that possibly the words were more familiar to some individuals than to others.

8. Penmanship Tests

*The Holmes Test.** The Holmes Test for the Speed of Writing is given as follows: A short sentence† is written on the board by the examiner and copied by the children at the top of their papers. After the directions for the test have been given, the sentence is read in concert by the class several times, so that it is fairly well learned before the test begins, and little or no time is lost by the pupils because of forgotten copy.

After two short practice periods of 15 seconds each, which are given for the purpose of "warming up," the sentence is written repeatedly for a one-minute and a four-minute period. The number of letters written per minute in each period is found, and the average of these is taken as the pupil's speed of writing.

---

*This test was arranged by Professor Henry W. Holmes and used with his kind permission. Records for comparison with the present results were also made available through his courtesy.

†The sentence is "Jolly kings bring gifts while happy maids dance." It will be noticed that this sentence contains most of the letters and the common letter combinations.

TABLE 81. SPEED OF HANDWRITING
Letters per Minute

Grade	2	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	38	66	75	100	112	111	111	1	1
Columbia	35	32	51	87	89	92	89	3	3.5
Kirksville	31	39	59	69	83	71	79	6	6
Maryville	39	54	51	55	83	89	83	5	5
Springfield	44	69	75	97	110	96	104	2	2
Warrensburg	45	44	60	79	82	92	87	4	3.5
All Schools	40	49	62	80	92	92	93		

Speed of Handwriting. Cape Giradeau easily leads in this test, with Springfield second. Columbia and Warrensburg stand so close together that they cannot be separated by any method of comparison.

In this test there seems to be regular progress through the grades. Three of the schools, however, show no advance after the sixth grade, and in two of these, Kirksville and Springfield, there is a falling off in the seventh grade.

Quality of Writing. The Holmes Test for the Quality of Writing provides for the rating on the Ayres Scale of three samples of the pupil's penmanship. The average of these three ratings is taken as the final score.

The first of the three samples rated was the paper written in the four-minute speed test; the second, a short story written by the pupils from dictation, and the third the "reproduction" paper from the reading test. In the latter case the pupils did not know that the paper was to be graded for penmanship. Each of the papers was rated independently by two of the examiners, and the two ratings were averaged for the score of the paper.

TABLE 82. QUALITY OF HANDWRITING
Scores on Ayres Scale

Grade	2	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	33	36	43	39	42	47	42	4	3
Columbia	50	42	46	43	44	49	47	2	2
Kirksville	35	32	39	36	38	35	37	6	6
Maryville	32	41	38	43	45	37	44	3	4.5
Springfield	34	29	33	36	43	40	41	5	4.5
Warrensburg	35	39	43	48	47	51	51	1	1
All Schools	35	36	39	41	42	45	44		

Warrensburg and Columbia stand very close in this test, but the former school has a little the better record in the upper grades. Cape Giradeau, Maryville, and Springfield are practically tied by the Upper Grade Median method of comparison, but the other methods give the advantage to the first school. Maryville would stand higher were it not for the poor record in the seventh grade. Kirksville does quite as well as most of the other schools in the lower grades, but makes no advance in the upper grades. The records of the upper grades are from one to two grades behind the other schools.

While the work of rating the papers was going on, it became evident to the examiners that there was a difference in the quality of the writing on the three different sets of papers. Separate medians were therefore found for each of the tests, and these are given in the table below.

TABLE 33. COMPARISON OF QUALITY OF HANDWRITING ON DIFFERENT TESTS

Grade	2	3	4	5	6	7
Medians in						
Dictation	35	35	40	43	45	50
Speed Test	33	35	35	38	38	43
Reproduction	35	35	40	42	45	45
Average of						
3 Tests	35	35	39	41	42	45

The differences are not very great up to the fourth grade, but thereafter the quality of writing in the Reproduction and the Dictation tests is slightly better than that in the Speed test. In the seventh grade the writing in the Dictation test is of better quality than in either of the other two tests. The quality in the Reproduction test is better than that in the Speed test. It is likely that the quality of writing in the Reproduction test is of most significance, because this was written in the pupils' ordinary handwriting.

The ranking of the different schools is practically the same in all three tests, as may be seen from the tables in the appendix.

For efficiency in writing, a proper balance must evidently be maintained between speed and quality; that is, neither speed nor quality of writing can be bettered at the expense of the other. When both factors are taken into consideration, the Warrensburg school stands first in the ranking. It has the best writing from the standpoint of quality combined with fair speed. Cape Giradeau and Columbia come next, with Springfield, Maryville, and Kirksville following in the order named.

4 Reading Tests

a. *The Kansas Silent Reading Tests.* The purpose of this test is to determine "the ability of pupils to get meaning from the printed

page." It consists of a series of "exercises," in the form of short paragraphs, in which the child is given certain directions to follow or in which he is required to solve a simple problem. Each exercise is so arranged that the response of the child is either right or wrong. A folder containing a set of these exercises is given to each child in the class, and all begin to work at a signal from the examiner. The score is determined by the number of exercises which are completed in five minutes, the weighting of each exercise in the final score having been calculated by the author of the tests.

The median scores obtained in this test are recorded in the following table. The slight decrease in the scores of the sixth grade is due to the fact that two sets of exercises are used, one set for the third, fourth, and fifth grades, and another for the sixth, seventh, and eighth.

TABLE 34. KANSAS SILENT READING TESTS

	Test I			Test II		6 & 7 combined		
Grade	3	4	5	6	7	Score	Rank	Final Rank
Cape Giradeau	18.1	16.3	23.8	10.6	30.0	19.7	1	1
Columbia	6.8	12.4	16.3	15.4	16.3	15.4	4	4
Kirksville	5.2	9.8	12.8	16.8	19.3	18.4	2	2
Maryville	7.2	8.5	14.6	11.7	15.8	12.8	6	6
Springfield	10.5	13.4	18.3	13.5	17.3	13.9	5	5
Warrensburg	10.3	10.3	12.5	12.9	20.1	16.7	3	3
All Schools	8.5	12.4	16.3	13.2	20.0	16.2		

In this test the ranking seems to be the same by all methods. Cape Giradeau stands first; but it should be noted that the pupils in this school had taken this test several times during the year. The difference between any two successive schools is not great, although a fairly large gap exists between the lowest and the highest scores.

As may be seen in the corresponding table in the appendix, the reliability of some of the scores in the third and fourth grades is rather questionable because of their high variability.

b. *Holmes Reading Tests.** In the first of these tests the pupils are given slips of paper on which is printed a simple story, and a period of 20 seconds is allowed for reading. When the signal to stop is given, each child makes a mark under the word which he is reading, and then finishes the reading of the story. When all are done, the class is asked to reproduce the story in writing. Finally, a set of questions on the main points of the story is given, and the pupils are asked to write the answers.

*These tests were arranged by Professor Henry W. Holmes and used with his kind permission.

The speed of silent reading is scored in the number of words read per minute; the reproduction is scored according to the number of ideas which are correctly reproduced by the pupil, and the answers to the questions are marked on a percentage basis according to definite values assigned them.

Before giving this test a similar test was first given to familiarize the children with the mechanics of the test. The material was the same as that used for testing the silent reading in the Cleveland Survey. The schools are not ranked on the basis of this latter test, but the results are given in Table XX of the appendix.

TABLE 35. SPEED OF SILENT READING
Holmes Test

	Words per Minute						6 & 7 combined Score		Final
Grade	2	3	4	5	6	7		Rank	Rank
Cape Giradeau	147	384	237	222	234	463	321	2	2
Columbia	186	147	219	309	327	324	324	1	1
Kirksville	126	207	180	321	264	276	264	5	5
Maryville	180	180	405	363	363	291	309	3	3
Springfield	138	237	188	364	276	237	249	6	6
Warrensburg	141	180	147	299	249	309	276	4	4
All Schools	147	198	198	309	276	318	291		

Speed of Silent Reading. It should be said at the outset of the discussion of this test that many of the scores of the lower grades are unreliable. This is not only because the variations are large, but because the schools do not take the same rank in the practice and the final tests. The sixth and seventh grade medians, however, are not subject to this criticism, and it is on the basis of these scores that the comparisons are made.

Columbia stands first, taking this rank over Cape Giradeau because of a considerably superior performance in the sixth grade. The ranking of the other schools is indicated in the above table.

TABLE 36. REPRODUCTION OF PASSAGE READ
Holmes Test

Grade	Score in Per Cent. 2	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	14	66	60	63	63	69	69	1	1
Columbia	29	37	29	57	49	57	49	6	6
Kirksville	11	43	40	51	60	63	63	2	2
Maryville	41	29	43	51	57	57	57	3.5	3.5
Springfield	26	49	54	49	57	54	57	3.5	3.5
Warrensburg	14	29	43	51	51	57	54	5	5
All Schools	20	37	43	51	54	60	57		

Reproduction Tests. In the reproduction of the passage read Cape Giradeau ranks first. It is noticeable that this school makes very little progress in this test, as the median of the third grade is nearly as high as that of the seventh. The ranking of the other schools may be seen in the accompanying table. None of the methods of comparison gave sufficient evidence for breaking the tie between Maryville and Springfield.

TABLE 37. ANSWERS TO QUESTIONS ON PASSAGE READ
Holmes Test

Grade	Score in Per Cent. 2	3	4	5	6	7	6 & 7 combined Score	Rank	Final Rank
Cape Giradeau	39	71	77	73	77	85	82	1	1
Columbia	53	53	44	68	70	73	70	4	3.5
Kirksville	19	39	49	64	69	68	69	5	5.5
Maryville	44	49	54	69	72	69	72	3	3.5
Springfield	38	59	68	77	78	77	77	2	2
Warrensburg	19	44	44	72	68	68	68	6	5.5
All Schools	37	50	52	69	72	72	73		

In the answers to questions on the passage read Cape Giradeau again takes first place, with Springfield second. The other four schools rank very close to one another, but the supplementary methods place Columbia and Maryville above the other two schools.

5. Composition Test

For this test the pupils of the seventh and eighth grades were asked to write a composition on any one of three following subjects: "The Most Exciting Day of My Life," "An Animal I Have Known," and "Something I Have Made." These topics were such that the pupil had a rather wide range of choice, while the product was kept fairly uniform.

The compositions were graded by the Harvard-Newton Composition Scale.

TABLE 38. COMPOSITIONS

School	Grade 7	8	Final Rank
Cape Giradeau	74	75	1
Columbia	70	..	3
Kirksville	64	72	4.5
Maryville	67	69	4.5
Springfield	75	75	2
Warrensburg	57	65	6
All School	67	70	

The comparisons in this test are based on the scores of the seventh grades alone because there is no eighth grade in Columbia, and the scores for the eighth grades at Kirksville and Maryville are based on so few cases as to be unreliable.

The medians show that there is little difference between Cape Girardeau and Springfield, but the distributions seem to indicate that the former school should have first place in the ranking. Columbia comes third, and is really a little more in advance of Maryville than the medians show. Kirksville and Maryville are so close that they are given the same rank. Warrensburg comes last, and is about a grade behind the other schools.

6. Final Ranking of the Schools

By way of summarizing the inter-school comparisons a final ranking of the schools on the basis of the results in all the tests has been made. Such a ranking, it is realized, must be made largely on the judgment of the investigators, and is likely to cover up more than it shows; as is apt to be the case when combinations of results are made. For the schools, the teachers, and the supervisory officers the most important things by far are the individual results; comparisons can be made only if great care is used in their interpretation.

The following tables give a summary of the findings presented in this section, together with the final ranking of the schools. The sign "&" in Table 39 indicates that two schools are tied in rank.

TABLE 39. SUMMARY OF RANKINGS IN ALL TESTS

		Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Courtis Tests	Addition—Speed	C. Giradeau	Maryville	Warrensburg	Columbia	Kirksville	& Springfield
	Addition—Accuracy	C. Giradeau	Springfield	Warrensburg	Columbia	Maryville	Kirksville
	Subtraction—Speed	Maryville	C. Giradeau	& Springfield	Warrensburg	Columbia	Kirksville
	Subtraction—Accuracy	Columbia	Warrensburg	Springfield	C. Giradeau	Maryville	Kirksville
	Multiplication—Speed	Maryville	C. Giradeau	Springfield	& Warrensburg	Kirksville	Columbia
	Multiplication—Accuracy	C. Giradeau	Springfield	Maryville	& Warrensburg	Columbia	Kirksville
	Division—Speed	Maryville	Springfield	Warrensburg	C. Giradeau	Columbia	Kirksville
	Division—Accuracy	C. Giradeau	Warrensburg	Springfield	Maryville	Columbia	Kirksville
	Stone Arithmetic Test	C. Giradeau	Springfield	Maryville	Columbia	Kirksville	& Warrensburg
Spell.	Ayres' List	Springfield	C. Giradeau	Maryville	Warrensburg	Columbia	Kirksville
	Boston List	Springfield	C. Giradeau	Warrensburg	Maryville	Columbia	Kirksville
Pen.	Speed	C. Giradeau	Springfield	Columbia	& Warrensburg	Maryville	Kirksville
	Quality	Warrensburg	Columbia	C. Giradeau	Springfield	& Maryville	Kirksville
Reading	Kansas Silent Reading	C. Giradeau	Kirksville	Warrensburg	Columbia	Springfield	Maryville
	Speed—Holmes	Columbia	C. Giradeau	Maryville	Warrensburg	Kirksville	Springfield
	Reproduction—Holmes	C. Giradeau	Kirksville	Maryville	& Springfield	Warrensburg	Columbia
	Questions—Holmes	C. Giradeau	Springfield	Columbia	& Maryville	Kirksville	& Warrensburg
	Composition	C. Giradeau	Springfield	Columbia	Kirksville	& Maryville	Warrensburg

TABLE 40. SUMMARY OF RANKS GIVEN EACH SCHOOL

	Courtis Tests									Spell.		Pen.			Reading					
	Addition—Speed	Addition—Accuracy	Subtraction—Speed	Subtraction—Accuracy	Multiplication—Speed	Multiplication—Accuracy	Division—Speed	Division—Accuracy	Stone Test	Ayres' List	Boston List	Speed	Quality	Kansas	Speed—Holmes	Reproduction—Holmes	Questions—Holmes	Compositions	Total of Ranks	Final Standing
Cape Giradeau	1	1	2.5	4	2	1	4	1	1	2	2	1	3	1	2	1	1	1	31.5	1
Columbia	4	4	5	1	6	5	5	5	4	5	5	3.5	2	4	1	6	3.5	3	72	5
Kirksville	5.5	6	6	6	5	6	6	6	5.5	6	6	6	6	2	5	2	5.5	4.5	95	6
Maryville	2	5	1	5	1	3.5	1	4	3	3	4	5	4.5	6	3	3.5	3.5	4.5	62.5	3
Springfield	5.5	2	2.5	3	3.5	2	2	3	2	1	1	2	4.5	5	6	3.5	2	2	52.5	2
Warrensburg	3	3	4	2	3.5	3.5	3	2	5.5	4	3	3.5	1	3	4	5	5.5	6	64.5	4

The tables show that Cape Giradeau easily ranks first. This school stands either first or second in all but three of the tests. Springfield stands second, while Maryville and Warrensburg are about tied for third place, with the advantage slightly in favor of the former. Columbia comes fifth, and Kirksville last—considerably below the other schools.

The relative standing of the schools is not, however, the most significant thing that the results of the tests show. Of much greater importance is the fact of the wide variation between these schools. Although there are, in practically every test, two or more schools in which the attainments are very close, yet in every test the difference between the highest and the lowest score is considerable. It does not seem that such variations should exist between schools which are engaged in similar work. Since it is unlikely that the selection of pupils or their general intelligence is sufficiently different in the various schools to produce such results, the explanation must be looked for in differences in the quality and amount of school training.

§3. COMPARISON OF RESULTS IN MISSOURI WITH RESULTS IN OTHER SCHOOL SYSTEMS

But few of the tests given above have been tried out in schools of sufficiently different character, location, and social condition to furnish a basis for reliable comparisons between schools which differ in these respects. Comparisons must, therefore, be made with the utmost caution. They should, moreover, be interpreted for the most part by teachers and supervisory officials who are in close touch with prevailing conditions.

As pointed out earlier in this report, recent developments in group methods for the measurement of intelligence have made it possible to investigate the native abilities of the pupils in various classes, schools and communities with results which force us to conclude that these abilities vary to considerable extent. It is, therefore, probably no more valid to compare the scores of a group of pupils in arithmetic without any consideration of their general mental levels than it would be to compare them without any reference to their attained school grade. It has been shown already that considerable evidence exists, especially in the age distributions, for supposing fundamental differences between the quality of the pupils in the various Missouri schools. It is altogether probable that like differences exist between the Missouri pupils and those from whose records the several standards were obtained.

In comparing the results in the schools studied above with results obtained elsewhere, it must, therefore, be kept in mind that the comparisons are made in many instances between schools of very different character and condition.

1. Arithmetic Tests

Courtis Tests. The standards chosen for comparison in this test are, in the first place, the latest* standards published by Courtis. These standards were obtained by giving the tests to several thousand pupils in all parts of the country. They include, however, very few results from Missouri or neighboring states.

The second set of standards consists of the results in three public systems of the state. These are more valuable than the above standards for the comparisons because it may be assumed that they were obtained from a somewhat similar school population.

As the results are much the same in all four operations, detailed comparisons will be made only for the results of the test in addition. The reader may easily make comparisons for himself of the results of the other tests by a study of the comprehensive tables in the appendix.

---

*At the time when this report was originally prepared.

TABLE 41. SPEED OF ADDITION IN COURTIS TESTS

Grade	3	4	5	6	7	8*
Training Schools	3	4	6	6	7	8
Courtis Standards	4	6	8	10	11	12
Columbia, Mo.	..	5	6	7	8	..
Piedmont, Mo.	3	6	5	6	6	7
Excelsior Springs, Mo.	..	6	6	6	7	8

Speed of Addition. The training schools are, throughout, from one to three grades behind the Courtis standards. Except in the third and fourth grades no one of the schools equals or surpasses these standards. On the other hand, some of the scores compare favorably with those obtained in the three public schools. In the three upper grades there is little difference.

TABLE 42. ACCURACY OF ADDITION IN COURTIS TESTS

Grade	3	4	5	6	7	8
Training Schools	0%	20%	33%	33%	57%	50%
Courtis Standards	41	64	70	73	75	76
Columbia, Mo.	..	60	50	43	63	..
Piedmont, Mo.	25	50	40	50	50	57
Excelsior Springs, Mo.	..	33	33	67	57	63

Accuracy of Addition. In this test the training schools are from two to four and a half grades behind the Courtis standards. Only four grades out of the whole number, three in Maryville, and one in Cape Giradeau, are approximately equal to the standard scores. Compared with the public schools the training schools fare little better in the lower grades, but in the seventh grade the results are about equal.

As has already been stated, the results in the other operations are all practically the same as those in addition. The training school scores are uniformly from one to five grades behind the Courtis standards, with only a single grade here and there equalling or approximating the latter. Comparisons with the scores from the Missouri public schools show that the training school scores are about equal to the lowest medians in the former group.

The joint performance of grades 4 to 8 inclusive of the training schools in the Courtis Tests in comparison with the Courtis Standards is shown graphically in Chart B on p. 46. The extreme of variation in the schools in grades 6 and 7 is also shown in the chart.

*The results of the eighth grade are, of course, incomplete, as they do not include any grade VIII pupils from Columbia, and only part of those from Kirksville and Maryville.

TABLE 43. STONE REASONING TEST

Grade	3	4	5	6	7	8
Training Schools	1.0	1.6	3.0	4.0	6.4	7.0
Starch Standards	4.5	6.2	7.8	9.4	11.0	12.6
Butte, Mont.	..	..	2.2	3.9	5.8	7.7
Laporte, Ind.	..	..	3.4	4.6	8.1	8.6
Salt Lake City, Utah	..	..	3.7	6.4	8.6	10.5
Boston, Mass., 1916	..	..	..	4.4	5.6	7.6
Brookline, Mass., 1916	..	..	4.0	6.2	..	..
Fall River, Mass., 1917	..	..	..	5.0	7.6	..

Stone Reasoning Test. The standards seem to fall into two groups, and the scores in one of these is considerably higher than those in the other. The training school medians are very close to the lower set of scores, running in some grades a little above, and in others, a little below. In none of the schools do the scores equal those of the higher group.

2. Spelling Tests

TABLE 44. AYRES' SPELLING TEST

Grade	2	3	4	5	6	7	8
Training Schools	81%	74%	64%	78%	74%	73%	77%
Ayres Standard	79	79	79	79	79	79	79
Brookline, Mass.	..	..	84	87	89	90	86
Fall River, Mass.	..	..	..	..	86	89	..

Words from Ayres' List. The words were so chosen that the standard score in each grade was 79%. The results in the training schools are only slightly below this standard, except in the fourth grade, where the average is about a grade below.

The Brookline scores are approximately a grade ahead of both the training school and the standard averages, and the same may be said of the Fall River averages.

Many of the grades in Cape Giradeau, Maryville, and Springfield show results which are higher than the Ayres standards, and which are approximately equal to the Brookline averages. It is interesting to note that in Columbia, where little attention is given to formal spelling work, the Ayres standard is practically attained in the seventh grade. Some of the results at Warrensburg and Kirksville are especially low.

TABLE 45. BOSTON SPELLING WORDS

Grade	2	3	4	5	6	7	8
Training Schools	57%	50%	42%	58%	55%	45%	57%
Brookline, Mass.	..	..	61	69	84	66	80
Fall River, Mass	..	..	..	..	72	60	..

Words from Boston List. The results with the Boston list are not as satisfactory as with the Ayres list, but this was to be expected because the words are harder and not so familiar to the pupils. The Boston scores with these words run from 86 to 95 per cent., but this is only after the words have been studied for one or two lessons. The question arises: Would a period or two of training on the Boston list of words in these schools raise the average so as to approximate the Boston standards? This is an experiment that the schools might well make.

The training school scores are also decidedly below those of Brookline where the conditions of testing were the same. Among the training schools there is none which even approximates the standards. Maryville is the only school in which any of the averages are over 70 per cent.

3. Penmanship Tests

TABLE 46. SPEED OF WRITING

Grade	2	3	4	5	6	7	8
Training Schools	40	49	62	80	92	92	102
Newton, Mass.	39	55	59	73	85	94	102
Brookline, Mass.	..	..	..	76	87	90	98
Fall River, Mass.	..	..	..	..	86	89	..

Speed of Writing. The training school medians are very close to those of the schools chosen for comparison. In the fifth and sixth grades they surpass slightly these other schools. There are many cases of low scores in the individual schools, but these are for the most part in the lower grades, which matters little, since we are concerned chiefly with the final product of the schools. The only exception to this is a low score in the seventh grade at Kirksville, which is about two grades behind the scores attained elsewhere.

TABLE 47. QUALITY OF WRITING

Grade	2	3	4	5	6	7	8
Training Schools	35	36	39	41	42	45	47
Newton, Mass.	54	50	45	48	51	50	53
Brookline, Mass.	..	..	..	44	46	47	49
Starch Standards*	27	33	37	43	47	53	57
South Bend Standards*	..	40	40	50	50	60	60
Cleveland, Ohio*	..	..	..	45	48	50	55
Fall River, Mass.	..	..	..	..	44	47	..

*The test given in this case was not the Holmes Test, but the writing was rated on the Ayres Scale.

Quality of Writing. The training school medians are, after the fourth grade, from one to two grades behind the standards. There is not much variation between the schools except for the fact that Kirksville stands far below the rest, making no advance from grade to grade, and not surpassing the second grade median even in the seventh grade.

4. Reading Tests

TABLE 48. KANSAS SILENT READING TESTS

Grade	3	4	5	6	7	8
Training Schools	8.5	12.5	16.3	13.2	20.0	18.4
Standards	6.0	9.9	13.7	13.4	16.5	18.8

Kansas Reading Tests. The standards for this test are the medians obtained from the results of giving the test to about 9000 pupils in 19 different cities. The training school medians are well above the standards except in the sixth and eighth grades, where the two are equal. The superiority of the training schools amounts to about two grades in the seventh and about one grade in the third, fourth, and fifth grades.

Cape Giradeau is in a class by itself in this test, but this is doubtless due to practice, as the test had been given previously in this school. Since the high scores in this school might influence the medians of the combined grades, these were found for the other five schools alone. The only appreciable difference is in the combined eighth grade median, which fell to 15.0, which is approximately equal to the seventh grade standard.

In the other schools there are very few grades in which the medians do not practically equal the standards. The chief exception to this is the seventh grade, in which the training school scores are all high. There are also several high medians in the Springfield grades; whereas in some of the grades at Kirksville and Maryville the scores are low.

TABLE 49. SPEED OF SILENT READING

Grade	2	3	4	5	6	7	8
Training Schools	147	198	198	309	276	318	363
Holmes Standards	...	...	...	249	276	309	309
Brookline, Mass.	...	...	...	249	283	309	309
Fall River, Mass.	...	...	...	...	270	309	...

Speed of Silent Reading. In this test the record of the training schools is a little better than the standards. The combined fifth grade median seems extraordinarily high, and is probably due to some factor of selection, which is not apparent.

There is no one school which is uniformly above or uniformly below the standards in all grades, but each school has some high and some low scores. In Kirksville the low medians are found in the upper grades.

TABLE 50. QUALITY OF REPRODUCTION OF PASSAGE READ

Grade	2	3	4	5	6	7	8
Training Schools	20%	37%	43%	51%	54%	60%	60%
Holmes Standards	..	..	..	54	54	60	60
Brookline, Mass.	..	..	..	57	57	57	60
Fall River, Mass.	..	..	..	..	54	57	..

Reproduction Tests. In the quality of reproduction of passages read there is little difference between the training school medians and the standards. Such differences as there are come in the fifth grade.

Only a few of the grade scores in the various schools are far removed from the standard medians. In Cape Giradeau the scores are consistently above the standards; in every grade after the second they equal or surpass the eighth grade standard.

TABLE 51. ANSWERS TO QUESTIONS ON PASSAGE READ

Grade	2	3	4	5	6	7	8
Training Schools	37%	50%	52%	69%	72%	72%	75%
Holmes Standards	..	..	..	63	69	72	73
Brookline, Mass.	..	..	..	62	64	68	73
Fall River, Mass.	..	..	..	..	66	69	..

In the answers to questions about the passages read, the training school medians are again not very different from the Holmes standards. They are from a grade to two grades ahead of the results in the Brookline and Fall River schools. As in the preceding test, the Cape Giradeau scores are very high from the third grade on.

5. Composition Test

TABLE 52. COMPOSITION TEST

Grade	7	8
Training Schools	67	70
Newton, Mass.	..	75
Bloomington, Ind.	61	67
Port Townsend, Wash.	53	58
Brookline, Mass.	..	70
Fall River, Mass.	65	..

In the composition test for the seventh grade, the training school medians are considerably above those of the schools with which they are compared, with the exception of Fall River. The difference is not so large in the eighth grade, but it must be remembered that the Missouri figures in this grade are not complete.

The differences between the medians in the separate schools and the standards follow in the main the same course as the combined grade medians. The scores in Cape Giradeau, Columbia, and Springfield are especially high.

6. Summary

On the whole, it may be said that the training schools compare not unfavorably with the other schools in which records are available. At least, it may be said that there are some standards in each set of results which the training school medians equal or surpass. The greatest discrepancies come in the results of the Courtis tests, but, although the training school medians are from one to five grades below the Courtis Standards, they are very close to the results obtained in three Missouri public schools. The general tendency seems to be for the training school combined grade medians to fall about half a grade below the standards.

Among the various schools, the Cape Giradeau and Springfield medians tend to surpass the standards, while the results from Kirksville are almost uniformly below. This is another evidence of rather wide variation among the schools.

§4. PROGRESS IN THE SCHOOLS

METHODS OF STUDY

Attention was called in Section 1 to the fact that the efficiency of a school may be judged, in part, at least, by the advance in the various subjects which the pupils of the school make from grade to grade. This was one of the methods used above to estimate the rank of a school. As was pointed out in the previous discussion, this method may lead to false conclusions since we are not dealing with the progress of the same pupils through the grades, but with the comparative standing of different pupils in the various grades. For example, if we compare a relatively inferior second grade with a relatively superior eighth grade, the results attained in that school would appear to be greater than they really are, as would be apparent had we been able to trace the actual improvement of this second grade from year to year up to the eighth grade. Or the opposite comparison may be made, namely, between a relatively superior second grade and a relatively inferior eighth grade. In this case the school would not be given credit for the improvement which would probably have been found, had we been able to follow the progress of the same individual pupils from this second grade until they reached the eighth grade.

It is to be noted, however, that when comparisons are made in many different subjects of study the relatively inferior and the relatively superior grades are likely to be discovered, since the high or the low scores are generally maintained consistently in all the various subjects of study. By such observations as these and by other like considerations, mentioned in Section 1, e.g., the comparative ages of the pupils in question, it is possible to make valid comparisons. It should, however, always be kept in mind that the progress discussed below is not that of the same pupils, but that of the school as a whole. The records simply tell how much superior the standing of the eighth grade pupils is to that of the second grade pupils in the same school. If the average second grade pupil writes at the rate of 39 letters per minute and the average eighth grade pupil writes at the rate of 102 letters per minute the amount of progress in the rate of writing in this school is 63 letters per minute.

An important element in the comparison is the uniformity or regularity in the progress from grade to grade. There may be a regular advance in each grade, or all the gains may be made in one or two grades with little or no gain in the other grades. Grading of pupils and the outlining of courses of study is done, for the most part, in the belief that regular and uniform progress is the rule. It will be interesting to see in how far the facts brought to light in this study justify this belief.

The facts in regard to school progress are shown more clearly in the accompanying diagrams. Tables of numerical results are added in the appendix. It was not considered desirable to make diagrams for all the tests; the following selection was made:

1. Courtis Test, Speed of Addition.
2. Courtis Test, Accuracy of Addition.
3. Stone Reasoning Test.
4. Speed of Penmanship.
5. Quality of Penmanship.
6. Speed of Silent Reading.
7. Quality of Reproduction of Passage Read.

GRADE PROGRESS IN COMBINED GROUP

The Progress lines for all the schools taken together are shown in the last diagram, Chart A. They indicate, for the most part, a steady and gradual gain in the average attainments of the pupils from the second to the eighth grade. In many cases the lines do not rise as rapidly after the sixth grades as they do before this grade. This is what we should expect, because intensive drill in the formal subjects is about finished by the end of the sixth grade.

There is one noticeable exception to the general regularity of improvement. In several of the tests the scores of the fifth grade are as high as or higher than the scores of the sixth grade. This is especially noticeable in the speed of silent reading, and in the speed of addition. There is a slight improvement in the accuracy of addition.

GRADE PROGRESS IN THE SEPARATE SCHOOLS

1. *Cape Girardeau.* The third diagram from the left side of page shows the lines of progress for Cape Girardeau. The first four of these lines, representing the work in arithmetic and the speed of writing, show fairly steady progress through the grades. In the quality of the writing there is a decided rise in the fourth grade, but otherwise the progress is uniform. The speed of reading, however, gives a very odd line of progress. The third grade score is very high, followed by a decided drop in the fourth, fifth and sixth grades, and a marked gain in the seventh grade. The progress in the reproduction test is unlike that in any of the other tests. There is again a sudden rise in the third grade, followed by a plateau which is broken only by a slight rise in the seventh grade.

The high scores in the third grade may be due to the fact that the number of pupils in this grade is very small (only eight in most of the tests) and that they are possibly a select group. The low scores of the

CHART A. PROGRESS OF TRAINING SCHOOLS FROM GRADE TO GRADE IN CERTAIN SELECTED ABILITIES

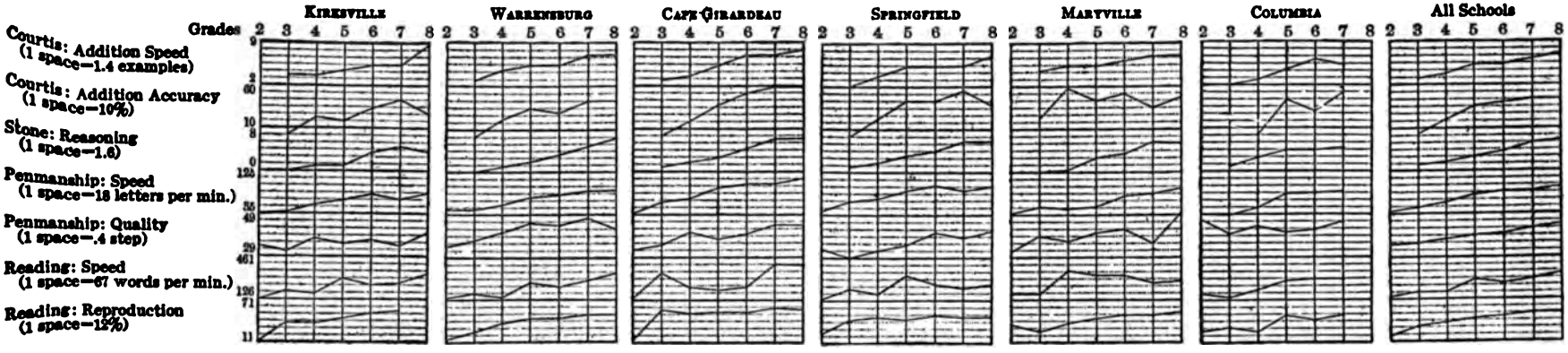
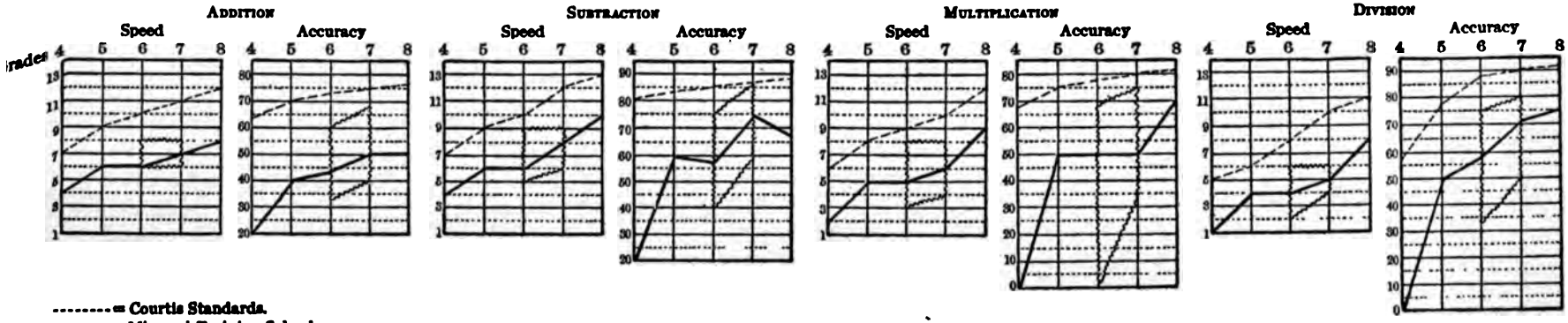


CHART B. JOINT PEFORMANCE OF TRAINING SCHOOLS IN COURTIS TESTS WITH EXTREME VARIATIONS IN GRADES SIX AND SEVEN



-------- = Courtis Standards.
——— = Missouri Training Schools.
~~~~~ = Extreme Variation, Grades 6 and 7.
~~~~~

fourth, fifth, and sixth grades in the speed of silent reading, may be counterbalanced in part by the high scores in these same grades in the quality of reproduction of the passage read.

When compared with the progress lines of the whole group the Cape Giradeau lines are for the most part parallel to the former and somewhat above them. This is especially true in the later grades, which shows that this school makes more than the average progress. Even in the speed of silent reading where the Cape Giradeau medians are below in the fifth and sixth grades, the seventh and eighth grade medians are considerably above those of the combined group.

2. *Columbia.* The progress in Columbia is shown in the next to the last diagram of Chart A. The most striking feature in this school is the plateau which occurs from the fifth grade on in nearly every test. This may be explained variously. It may be caused by the fact that little attention is given to the formal side of the studies in this school, especially in the later grades, or the seventh grade may be a relatively inferior grade. As there is no eighth grade in this school, it is not possible to get further evidence on this point. It should be said in passing that if the second explanation is true, this fact has caused Columbia to suffer in the comparisons with the other schools.

Another noticeable fact is that the quality of writing is high throughout all the grades of this school. When compared with the lines of the combined group the Columbia results are seen to be, in general, slightly lower, with the exception of the fifth grade; but the record in the quality of penmanship is everywhere above that of the combined group.

3. *Kirksville.* In the first diagram of the chart showing the grade progress in Kirksville there is no uniformity in the advances in the various subjects and no general tendency which can be observed. Regularity of progress is shown to some extent in the speed of addition, in the speed of writing through the sixth grade, and in the quality of reproduction. The other lines are very irregular. In comparison with the combined group the results at Kirksville are, with the exception of the record in the quality of reproduction, uniformly lower. This is especially true of the upper grades.

4. *Maryville.* In the diagram for the grades at Maryville, the lack of uniformity in progress is again the main feature. The line representing the Stone test shows the most regular progress; there is also considerable regularity in the gains made in the speed of addition. In the speed of silent reading, the line slopes downward from the fourth grade on, and in the quality of reproduction there is a marked plateau in the progress of the last four grades. Because of this irregularity, comparison with the combined group is difficult.

5. *Springfield.* The Springfield diagram brings out clearly the fact to which attention has been called previously, namely, that the seventh grade in this school seems to be a poor one. In five out of the seven lines there is no rise between the sixth and seventh grades, and in most cases there is a noticeable drop. Up to the sixth grade, the progress is in most cases regular. In the speed of reading, the line shows the most irregularity, and it is especially noticeable for the downward trend from the fifth grade on. In the quality of reproduction there is a plateau from the third grade on to the eighth.

A general statement concerning the relation of the Springfield lines to those representing the combined schools can hardly be made. In both the arithmetic tests and in the speed of writing, the Springfield records are slightly above the combined record, whereas in the quality of writing the combined record is higher, except in the sixth grade. In both the speed of silent reading and the quality of reproduction the tendency is for the school to be somewhat above in the earlier grades and somewhat below in the later grades.

On the whole, it may be said that the progress in this school is satisfactory, if the influence of what seems to be a relatively inferior seventh grade is discounted.

6. *Warrensburg.* The Warrensburg lines of progress indicate the most satisfactory state of affairs that is found in any of the schools. Progress seems to be steady and regular in all the tests. Only a few isolated scores break the regularities of the lines, and these are not important. In nearly all cases the records are very close to those of the combined group. The only exceptions are in the case of the Stone test, which is a little below, and in the quality of writing, which is a little above the standard.

SUMMARY

The study of the progress in the schools again brings out the fact that there are many inter-school differences, as well as inequalities in the same school and grade in the standing of the pupils in the several subjects of study. Although there are many cases of regular and continued progress through the grades, the ups and downs in the records show altogether too frequently that grade standards are lacking. Many of the irregularities are doubtless to be explained by the limitation pointed out at the beginning of this section, namely, that the progress has not been measured by the attainments of the same groups of pupils in successive years; but the fact still remains that, in many instances, grade position gives litttle or no indication of the character of work or the attainments of the pupils.

§5. CORRELATIONS

METHODS

The question of the relative standings of pupils in the various tests raises an interesting set of problems. A few of the possible correlations having practical as well as theoretical interest have been studied below. Most of the correlations were calculated by the Spearman "foot-rule" formula and then translated into the equivalent co-efficients of correlation of the Pearson method.* Since there is some question concerning the reliability of this method of calculation, a number of the correlations were also obtained by the usual Pearson formula as a check on the results. As will be seen below, the results by the two methods were in sufficient agreement to justify the use of the shorter method alone in the remaining correlations. The various studies are reported in the tables which follow.

TABLE 58. CORRELATION BETWEEN THE SPEED AND QUALITY OF WRITING

	Number of Cases	r	P.E.
Cape Giradeau	107	.60*	.04
Columbia	74	.12†	.07
Kirksville	97	.48	.05
Maryville	72	.26	.07
Springfield	99	.44	.05
Warrensburg	161	.36	.05
All Schools	610	.49	.02

RESULTS

The Relation of Speed to Quality of Writing. In all the schools, with the exception of Cape Giradeau, the correlation between the speed and the quality of writing is low. In Columbia there is little or no relationship, and the correlation is also especially low at Maryville. This would seem to indicate a lack in the co-ordinate development of these two phases of penmanship.

---

*By means of Table 37, p. 169 of Thorndike's "Mental and Social Measurements." The formula is $r = \sin(\sigma/2)R$, in which "r" stands for the Pearson co-efficient and "R" for the Spearman.

*In this case the Pearson Co-efficient was .60.

†In this case the Pearson Co-efficient was .17.

TABLE 54. CORRELATION BETWEEN QUALITY OF WRITING IN SPEED TEST AND QUALITY OF WRITING IN REPRODUCTION TEST

	Number of Cases	r	P.E.
Cape Giradeau	104	.63*	.04
Columbia	73	.73	.03
Kirksville	92	.63†	.04
Mayrville	72	.70‡	.04
Springfield	98	.78	.03
Warrensburg	164	.69§	.03
All Schools	603	.72	.01

The Quality of Writing in Two Different Tests. In one of these tests the pupils knew that their penmanship was being tested and would be graded; in the other, the reproduction test, they did not suspect that the quality of writing would be especially considered. The degree of correlation, as shown in the above table, may, therefore, indicate the extent to which the instruction in penmanship has "carried over" to the ordinary school work of the pupils. The inter-school differences are, in this case, unimportant.

TABLE 55. CORRELATION BETWEEN THE SPEED OF SILENT READING AND THE QUALITY OF REPRODUCTION OF PASSAGE READ

	Number of Cases	r	P.E.
Cape Giradeau	104	.35	.06
Columbia	75	.37	.07
Kirksville	84	.52	.05
Maryville	71	.22	.07
Springfield	100	.23	.06
Warrensburg	159	.44	.04
All Schools	593	.39¶	.02

The Speed of Reading and the Quality of Reproduction. The correlations between the speed of reading and the ability to reproduce accurately what is read are, in general, of small amount, and are especially so in Maryville and Springfield. The highest correlation is attained at Kirksville. Since the correlations, although positive, are of small amount, they do not altogether bear out the statement commonly

*In this case the Pearson Co-efficient was .61.
†In this case the Pearson Co-efficient was .58.
‡In this case the Pearson Co-efficient was .62.
§In this case the Pearson Co-efficient was .74.
¶The Pearson Co-efficient is also in this case .39.

made that the fast readers also retain or reproduce best what they have read. Most of the studies bearing on the matter have been made on adults, and the question has not been sufficiently studied in the case of children. The differences found may well be in part the result of differences in the methods of school training.

TABLE 56. CORRELATION BETWEEN THE QUALITY OF REPRODUCTION AND THE ANSWERS TO QUESTIONS IN THE HOLMES READING TEST

	Number of Cases	r	P.E.
Cape Giradeau	106	.70	.03
Columbia	75	.76	.03
Kirksville	82	.77	.03
Maryville	70	.55	.05
Springfield	102	.48	.01
Warrensburg	158	.68	.03
All Schools	593	.70	.01

Quality of Reproduction in Different Tests. The correlation studied in table 56 is one in which a very high co-efficient is to be expected, for the two tests are evidently similar in character. The chief difference is that the questions cover the main points of the story, while the reproduction calls for all the details that can be remembered. The correlations are lowest in Maryville and Springfield. These are the same schools in which there was little relation between the speed of reading and the quality of reproduction. There is some lack of uniformity in the results from the several schools, although the differences are not large.

TABLE 57. CORRELATION BETWEEN THE REPRODUCTION OF A PASSAGE READ TO THE PUPILS AND A PASSAGE READ BY THEM

	Number of Cases	r	P.E.
Cape Giradeau	100	.75	.03
Columbia	72	.80	.03
Kirksville	31	.59	.08
Maryville	69	.53	.06
Springfield	96	.58	.05
Warrensburg	157	.56	.03
All Schools	525	.64	.01

"Auditory" and "Visual" Reproduction. The passage read to the pupils was very similar in form to the passage read by them. It was a short story, based upon thirty-five ideas, which was read to each

class by the teacher or her assistant. After it had been read through once, the pupils were asked to write what they could remember of the story in exactly the same way as was done in the case of the corresponding passage which they had read for themselves.

As may be seen in the above table, there is a fair amount of correlation in the standings of pupils in the two tests. The majority of those who do well in one test do well in the other. School training may have a very considerable effect on the result, and it is, therefore, worthy of note, in view of the different methods of training practiced at Columbia, that the highest correlation is found in that school.

§6. RECOMMENDATIONS AND SUGGESTIONS FOR FURTHER STUDY

The findings of this investigation will gain in value if they may be supplemented by the comments and interpretation of the teachers and supervisors who are conversant with the special characteristics, aims, problems, and limitations of the several schools, and if they may receive rigorous criticism and discussion. In this way only will the findings presented secure their proper evaluation.

The large differences between the accomplishments of the pupils in the various schools certainly deserve further attention. Some of these differences may be due to factors which have not been adequately canvassed in this study; but for the most part, they are to be attributed to differences in the standards and quality of instruction. Unless the mental calibre of the pupils in the various schools is very different, it should be possible for these schools to maintain approximately equal standards of scholarship. If certain schools are handicapped by the grade of pupils received for instruction, this fact should be ascertained and brought to the attention of the normal school students who are serving as practice teachers in these schools. Otherwise, the latter may easily be satisfied with lower standards in the schools and classes over which they are subsequently put in charge, to the detriment of instruction throughout the state.

One way of obviating this difficulty would be to arrange for an exchange of practice teachers between the state normal schools. Such a practice might well lead to greater uniformity of standards and to practical co-operation between these schools.

It may need again to be emphasized that the examiners are well aware that standing in these tests of the more formal activities of school children is not the sole criterion of school efficiency; but it is believed that when certain schools fall far below the attainments of other schools in the formal subjects, it may well be inquired whether or not these schools are giving compensating advantages. The causes of interschool differences may, therefore, be profitably investigated. If individual weakness or merit is exposed, it should not be for the purpose of drawing invidious distinctions, but in order to lead to helpful co-operation between the schools.

After such comparisons, each school should be free to work out its own standards. When once its standards have been set up, the tests are useful to determine whether or not the requirements are being met in the separate classes. In this connection it should be noted that if school tests are given too prominent a place there is danger that some teachers will be encouraged to over-stress the drill in the formal subjects

in order to keep the classes up to the standards. They will do this to the neglect of other subjects of study in which the results of instruction are not as tangible or as easily measured. This is a difficulty which can be met only by efficient supervision.

By the use of these tests it is possible to single out pupils who need individual attention. Even the best of teachers are not always cognizant of the limitations and needs of individual pupils. A cross-section of a pupil's abilities such as is given by a series of tests frequently reveals facts about him hitherto unrecognized.

When these school tests are supplemented, as they may well be in many individual cases, by "intelligence" tests, not only is the proper differentiation of backward or defective pupils for special class work or individual instruction made more certain, but pupils of superior endowment may not infrequently be recognized and given opportunities for more rapid advancement or specialization of instruction.

In such ways may the findings of this investigation be profitably extended.

APPENDIX

§1. COMPLETE TABLE OF RESULTS

In this section will be found comprehensive tables giving the complete results of the study.

Unless otherwise noted, the results are stated in medians which are the figures in the first line under each school. The second line of figures show the median deviations or probable errors. The figures in italics are the number of cases from which the above measures are calculated.

In the Courtis tests "A" stands for "attempts" and "R" for "rights."

TABLE I

COURTIS TESTS: SPEED OF ADDITION

Grade	3		4		5		6		7		8	
	A	R	A	R	A	R	A	R	A	R	A	R
Cape Giradeau	3	0	4	0	6	3	8	3	8	5	9	7
	1	0	1	2	1	1	2	2	2	2	2	2
	9		*12*		*15*		*13*		*21*		*15*	
Columbia	2	0	3	0	5	2	7	2	6	4	. .	.
	1	0	1	0	1	1	2	1	2	1	. .	.
	10		*18*		*11*		*12*		*14*			
Kirksville	4	0	4	1	5	1	6	2	6	3	10	1
	1	0	1	0	1	1	2	1	2	2	4	1
	16		*17*		*12*		*22*		*18*		*8*	
Maryville	5	1	6	4	6	4	7	5	8	3	8	5
	1	1	1	2	1	2	2	2	1	1	1	2
	10		*10*		*16*		*8*		*12*		*9*	
Springfield	2	0	4	1	6	3	6	4	6	4	8	3
	1	0	1	1	2	2	2	2	1	1	2	2
	14		*15*		*12*		*16*		*16*		*14*	
Warrensburg	3	0	5	1	6	3	6	2	8	4	8	4
	1	0	1	1	1	2	1	2	1	2	1	2
	20		*15*		*20*		*21*		*41*		*30*	
All Schools	3	0	4	1	6	2	6	2	7	4	8	4
	1	0	1	1	1	2	1	2	1	2	1	2
	79		*87*		*86*		*92*		*122*		*76*	
Courtis Standards*	4		6		8		10		11		12	
Columbia, Mo.	. .	.	5	3	6	3	7	3	8	5	. .	.
	. .	.	*142*		*145*		*124*		*119*			
Piedmont, Mo.	3	1	6	3	5	2	6	3	6	3	7	4
	50		*34*		*26*		*47*		*29*		*19*	
Excelsior Springs, Mo.			6	2	6	2	6	4	7	4	8	5

*Bulletin No. 4. Courtis Standard Research Tests, issued by The Department of Co-operative Research, 82 Eliot St., Detroit, Mich.

TABLE II

COURTIS TESTS: ADDITION ACCURACY

Grade	3	4	5	6	7	8
Cape Giradeau	0	20	43	60	67	67
	0	20	18	27	13	15
	9	*12*	*15*	*13*	*21*	*15*
Columbia	0	0	50	33	60	..
	0	0	27	19	22	..
	10	*18*	*11*	*12*	*14*	*0*
Kirksville	0	25	20	38	50	30
	0	11	20	22	25	30
	16	*17*	*12*	*22*	*18*	*8*
Maryville	25	67	50	60	40	55
	25	33	50	23	15	16
	10	*10*	*18*	*8*	*12*	*9*
Springfield	0	25	50	50	64	43
	0	25	25	17	19	14
	14	*15*	*12*	*16*	*16*	*14*
Warrensburg	0	25	38	33	50	50
	0	15	25	19	17	14
	20	*15*	*20*	*21*	*41*	*30*
All Schools	0	20	40	44	50	50
	0	20	23	23	17	20
	79	*87*	*86*	*92*	*122*	*76*
Courtis Standards	41	64	70	73	75	76
Columbia, Mo.	..	60	50	43	63	..
		142	*145*	*124*	*119*	..
Piedmont, Mo.	25	50	40	50	50	57
	50	*34*	*26*	*47*	*29*	*19*
Excelsior Springs, Mo.	..	33	33	67	57	63

TABLE III

COURTIS TESTS: SUBTRACTION

Grade	3		4		5		6		7		8	
	A	R	A	R	A	R	A	R	A	R	A	R
Cape Giradeau	3	0	4	2	8	5	8	5	8	7	9	6
	1	0	1	1	1	1	2	3	2	2	2	2
	9		*12*		*15*		*13*		*24*		*15*	
Columbia	1	0	1	0	4	2	6	3	7	6	.	.
	1	0	1	0	1	1	2	2	1	1	.	.
	10		*18*		*11*		*12*		*14*		*0*	
Kirksville	1	0	4	0	6	2	5	1	6	3	8	3
	1	0	1	0	3	2	1	1	2	2	2	1
	16		*17*		*12*		*22*		*18*		*8*	
Maryville	2	0	5	4	6	4	9	5	9	6	10	7
	1	0	2	1	1	2	2	2	1	2	1	1
	10		*10*		*16*		*8*		*12*		*9*	
Springfield	4	1	4	1	7	5	8	6	8	5	10	6
	1	1	1	1	0	2	2	2	2	2	2	2
	14		*15*		*12*		*16*		*16*		*14*	
Warrensburg	3	0	4	2	7	4	5	3	9	7	10	7
	1	0	1	1	1	1	1	1	2	2	2	2
	20		*15*		*20*		*21*		*41*		*30*	
All Schools	2	0	4	1	6	4	6	3	8	6	10	6
	1	0	1	1	1	1	2	2	2	2	2	3
	79		*87*		*86*		*92*		*122*		*76*	
Courtis Standards	5		7		9		11		12		13	
Columbia, Mo.	.	.	6	3	6	4	8	5	10	8	.	.
			142		*146*		*124*		*118*			
Piedmont, Mo.	2	0	4	2	5	2	6	3	6	3	8	6
	47		*35*		*25*		*48*		*29*		*19*	
Excelsior Springs, Mo.	.	.	6	2	7	3	8	5	9	5	10	7

TABLE IV

COURTIS TESTS: SUBTRACTION ACCURACY

Grade	3	4	5	6	7	8
Cape Giradeau	0	50	63	55	70	67
	0	25	20	33	13	18
	9	*12*	*15*	*13*	*24*	*15*
Columbia	0	0	50	43	86	..
	0	0	17	32	3	..
	10	*18*	*11*	*12*	*14*	*0*
Kirksville	0	0	33	40	60	50
	0	0	33	27	26	20
	16	*17*	*12*	*22*	*18*	*8*
Maryville	0	80	60	55	67	70
	0	30	26	18	21	8
	10	*10*	*16*	*8*	*12*	*9*
Springfield	25	20	71	75	67	60
	25	20	15	14	19	13
	14	*15*	*12*	*16*	*16*	*14*
Warrensburg	0	40	60	67	78	73
	0	11	18	13	13	16
	20	*15*	*20*	*21*	*41*	*30*
All Schools	0	20	60	57	75	67
	0	20	23	26	15	15
	79	*87*	*86*	*92*	*122*	*76*
Courtis Standards	49	80	83	85	86	87
Columbia, Mo.	..	50	67	63	80	..
	..	*142*	*146*	*124*	*118*	..
Piedmont, Mo.	0	50	40	50	50	75
	47	*35*	*25*	*48*	*29*	*19*
Excelsior Springs, Mo.	..	33	43	63	56	70

TABLE V

COURTIS TESTS: MULTIPLICATION

Grade	3		4		5		6		7		8	
	A	R	A	R	A	R	A	R	A	R	A	R
Cape Giradeau	0	0	3	0	5	3	8	4	7	4	10	6
	1	0	2	0	1	1	1	2	2	2	1	2
	9		*12*		*15*		*13*		*24*		*15*	
Columbia	"0"	"0"	1	0	1	0	3	0	4	2	. .	.
	0	0	0	0	0	0	1	0	1	2		
	10		*18*		*11*		*12*		*14*		*0*	
Kirksville	"0"	"0"	1	0	3	1	3	0	6	2	10	0
	0	0	1	0	2	1	1	0	1	1	3	0
	16		*17*		*12*		*22*		*18*		*8*	
Maryville	3	0	5	1	5	3	7	4	8	5	8	3
	1	0	1	2	1	1	2	1	1	2	2	2
	10		*10*		*16*		*8*		*12*		*9*	
Springfield	"0"	"0"	3	1	5	4	8	5	6	3	8	4
	0	0	1	1	1	1	2	1	1	1	2	2
	14		*15*		*12*		*16*		*16*		*14*	
Warrensburg	"0"	"0"	3	0	6	3	5	2	7	4	9	5
	0	0	1	0	1	2	2	2	1	2	2	2
	20		*15*		*20*		*21*		*41*		*30*	
All Schools	"0"	"0"	2	0	5	3	5	2	6	4	9	4
	0	0	2	0	1	2	2	2	2	2	2	2
	79		*87*		*86*		*92*		*122*		*76*	
Courtis Standards	0		6		8		9		10		11	
Columbia, Mo.	. .	.	5	2	5	3	7	4	10	7		
			139		*146*		*129*		*118*			
Piedmont, Mo.	2	0	4	1	5	2	5	3	6	4	8	5
	47		*31*		*24*		*47*		*29*		*19*	
Excelsior Springs, Mo.			5	2	5	2	6	4	7	5	9	5

TABLE VI

COURTIS TESTS: MULTIPLICATION ACCURACY

Grade	3	4	5	6	7	8
Cape Giradeau	0	0	67	63	75	55
	0	0	16	19	25	17
	9	*12*	*15*	*13*	*21*	*15*
Columbia	0	0	0	0	57	..
	0	0	0	0	29	
		18	*11*	*12*	*14*	*0*
Kirksville	0	0	25	0	33	0
	0	0	25	0	17	0
		17	*12*	*22*	*18*	*8*
Maryville	0	25	75	50	50	43
	0	25	25	17	12	10
	10	*10*	*16*	*8*	*12*	*9*
Springfield	0	0	80	67	57	45
	0	0	13	8	17	20
		15	*12*	*16*	*16*	*14*
Warrensburg	0	0	50	50	60	70
	0	0	17	36	20	20
		15	*20*	*21*	*41*	*30*
All Schools	0	0	50	50	50	50
	0	0	30	30	25	20
		87	*86*	*92*	*122*	*76*
Courtis Standards	..	67	75	78	80	81
Columbia, Mo.	..	40	60	57	70	..
		139	*146*	*129*	*118*	..
Piedmont, Mo.	0	25	40	60	67	63
	47	*31*	*24*	*47*	*29*	*19*
Excelsior Springs, Mo.	..	40	40	67	71	56

TABLE VII

COURTIS TESTS: DIVISION

Grade	3		4		5		6		7		8	
	A	R	A	R	A	R	A	R	A	R	A	R
Cape Giradeau	"0"	"0"	1	0	4	1	5	3	5	4	8	7
	0	0	1	0	1	1	3	2	2	2	4	4
	9		*12*		*15*		*13*		*21*		*15*	
Columbia	"0"	"0"	"0"	"0"	1	0	2	1	5	2	..	.
	0	0	0	0	0	0	1	1	2	1		
	10		*18*		*11*		*12*		*14*			
Kirksville	"0"	"0"	1	0	2	0	3	1	4	2	4	1
	0	0	1	0	1	0	2	1	1	1	4	1
	16		*17*		*12*		*22*		*18*		*8*	
Maryville	"0"	"0"	3	1	4	2	5	3	6	3	6	4
	0	0	2	1	1	1	3	2	1	1	2	1
	10		*10*		*16*		*8*		*12*		*9*	
Springfield	"0"	"0"	2	0	3	2	6	4	5	4	9	7
	0	0	0	0	1	1	1	2	1	1	1	2
	14		*15*		*12*		*16*		*16*		*14*	
Warrensburg	"0"	"0"	2	0	5	3	4	2	6	4	7	6
	0	0	1	0	2	2	1	1	2	2	2	2
	20		*15*		*20*		*21*		*41*		*30*	
All Schools	"0"	"0"	1	0	4	1	4	2	5	3	8	6
	0	0	1	0	1	1	2	2	2	2	3	2
	79		*87*		*86*		*92*		*122*		*76*	
Courtis Standards	0		4		6		8		10		11	
Columbia, Mo.	..		3	1	3	1	4	3	9	7		
			139		*144*		*124*		*119*			
Piedmont, Mo.	..		3	0	4	1	3	2	4	2	7	5
			23		*25*		*48*		*28*		*18*	
Excelsior Springs, Mo.	..		..		3	1	6	5	7	6	8	6

TABLE VIII

COURTIS TESTS: DIVISION ACCURACY

Grade	3	4	5	6	7	8
Cape Giradeau	0	0	33	75	80	83
	0	0	33	25	20	9
		12	*15*	*13*	*21*	*15*
Columbia	0	0	0	50	60	..
	0	0	0	50	20	
			11	*12*	*14*	*0*
Kirksville	0	0	0	33	50	44
	0	0	0	33	25	44
		17	*12*	*22*	*18*	*8*
Maryville	0	60	57	50	60	75
	0	40	23	25	20	15
		10	*16*	*8*	*12*	*9*
Springfield	0	0	67	71	75	80
	0	0	27	21	15	20
		15	*12*	*16*	*16*	*14*
Warrensburg	0	0	60	67	80	80
	0	0	20	33	20	20
		15	*20*	*21*	*41*	*30*
All Schools	0	0	50	57	71	75
	0	0	30	37	21	18
		87	*86*	*92*	*122*	*76*
Courtis Standards	..	57	77	87	90	91
Columbia, Mo.	..	33	33	75	78	..
		139	*114*	*124*	*119*	..
Piedmont, Mo.	..	0	25	67	50	71
		23	*25*	*48*	*28*	*18*
Excelsior Springs, Mo.	..	..	25	83	86	75

TABLE IX

STONE ARITHMETIC TESTS

Grade	3	4	5	6	7	8
Cape Giradeau	1.0	2.0	3.0	5.0	7.1	6.9
	1.0	1.0	1.0	2.0	1.0	3.1
	8	*13*	*15*	*13*	*22*	*16*
Columbia	1.0	2.5	4.6	4.6	5.0	...
	0.4	2.1	1.0	1.6	1.7	
	10	*18*	*11*	*14*	*11*	
Kirksville	0	1.0	1.0	4.0	5.2	4.0
	0	1.0	1.0	2.0	1.2	1.0
	15	*15*	*13*	*22*	*18*	*9*
Maryville	0	0.5	3.1	4.0	6.7	6.3
	0	0.5	2.0	1.2	0.7	1.3
	10	*10*	*16*	*8*	*12*	*10*
Springfield	1.0	2.0	3.6	4.6	6.8	6.8
	0.6	1.0	3.6	1.0	1.2	1.8
	16	*15*	*10*	*15*	*17*	*14*
Warrensburg	0	1.0	2.0	3.6	5.6	7.8
	0	1.0	1.0	1.8	2.0	1.4
	20	*15*	*20*	*21*	*41*	*32*
All Schools	1.0	1.6	3.0	4.0	6.4	7.0
	1.0	1.4	2.0	1.6	1.6	2.0
	79	*86*	*85*	*93*	*121*	*81*
Starch Standards*	4.5	6.2	7.8	9.4	11.0	12.6
Butte, Mont.	..	..	2.2	3.9	5.8	7.7
Laporte, Ind	..	..	3.4	4.6	8.1	8.6
Salt Lake City	..	..	3.7	6.4	8.6	10.5
			453	*468*	*409*	*284*
Boston, 1916	..	..	..	4.4	5.6	7.6
				1.6	1.8	2.4
				1243	*1128*	*1065*
Brookline, Mass.	..	..	4.0	6.2	..	..
			1.4	2.0	..	..
			322	*325*		

*These scores are based on results from 2515 pupils in 18 schools.

TABLE X

SPELLING ("AYRES'" LISTS)

Grade	2	3	4	5	6	7	8
Cape Giradeau	76	84	82	85	80	83	82
	14	8	13	12	10	13	12
	21	*8*	*11*	*14*	*13*	*22*	*14*
Columbia	not	60	57	66	67	76	no
	given	17	25	21	25	11	8th
		10	*19*	*11*	*12*	*14*	grade
Kirksville	67	66	53	73	64	58	57
	22	20	29	20	19	24	21
	17	*14*	*16*	*10*	*24*	*18*	*8*
Maryville	97	93	73	80	76	76	90
	2	8	17	18	17	20	12
	6	*10*	*9*	*16*	*8*	*12*	*10*
Springfield	95	85	72	89	81	79	76
	5	8	15	8	15	17	19
	16	*16*	*14*	*12*	*16*	*17*	*14*
Warrensburg	82	65	58	76	79	71	77
	15	17	22	20	17	20	19
	15	*21*	*16*	*20*	*21*	*41*	*30*
All Schools	81	74	64	78	74	73	77
	17	15	23	18	18	19	16
	75	*79*	*85*	*83*	*94*	*124*	*76*
Standard	79	79	79	79	79	79	79

The above scores are average percentages, and the figures in the second line are the average deviations. The figures in italics indicate the number of pupils tested in each instance.

The words for these lists were so chosen that the standard accuracy is 79%. That is, an average of 79% must be obtained to equal the standard.

In correcting the papers, credit was given for accepted forms of simplified spelling and for wrong words correctly spelled.

TABLE XI

SPELLING ("BOSTON" LIST)

Grade	2	3	4	5	6	7	8
Cape Giradeau	57	65	66	57	55	62	64
	14	14	16	4	27	21	20
	21	*8*	*11*	*14*	*13*	*22*	*14*
Columbia	not	38	28	35	53	36	no
	given	14	18	19	33	16	8th
		10	*19*	*11*	*12*	*14*	grade
Kirksville	36	33	29	59	39	33	41
	24	20	22	21	24	23	14
	17	*14*	*16*	*10*	*24*	*18*	*8*
Maryville	73	59	38	57	59	37	54
	20	25	20	21	26	20	22
	6	*10*	*9*	*16*	*8*	*12*	*10*
Springfield	73	61	71	71	74	50	62
	18	16	14	18	18	21	22
	16	*16*	*14*	*12*	*16*	*17*	*14*
Warrensburg	59	47	31	65	60	43	56
	17	24	18	28	24	24	24
	15	*21*	*16*	*20*	*21*	*41*	*30*
All Schools	57	50	42	58	55	45	57
	21	22	24	23	26	24	18
	75	*79*	*85*	*83*	*94*	*124*	*76*

The above scores are average percentages, and the figures in the second line are average deviations. The figures in italics indicate the number of pupils tested in each instance.

The standard accuracy for these words ranges from 86% to 95%. These results, however, were obtained only after the words had been studied for one or two lessons. None of the schools tested had used these words.

TABLE XII

HOLMES PENMANSHIP TEST

Speed of Writing—4-Minute Test

Grade	2	3	4	5	6	7	8
Cape Giradeau	34	60	62	97	106	109	121
	6	8	10	8	8	9	9
	21	*8*	*13*	*15*	*13*	*22*	*16*
Columbia	31	25	52	86	90	87	
	4	5	12	11	12	17	
	11	*10*	*18*	*11*	*11*	*14*	*0*
Kirksville	20	39	59	64	75	76	64
	10	12	8	7	13	21	19
	18	*15*	*16*	*12*	*21*	*18*	*7*
Maryville	35	48	48	54	72	85	99
	10	18	13	9	10	20	10
	6	*10*	*10*	*16*	*8*	*12*	*10*
Springfield	44	62	73	80	94	94	99
	7	8	8	13	7	8	8
	15	*16*	*14*	*10*	*15*	*17*	*14*
Warrensburg	46	38	58	76	80	88	93
	13	10	8	15	14	11	9
	13	*21*	*15*	*20*	*21*	*41*	*31*
All Schools	34	45	58	76	84	90	98
	8	15	11	15	15	18	14
	84	*80*	*86*	*84*	*89*	*124*	*78*
Cleveland*	..	..	..	60	70	76	80
Starch Standards*	31	38	47	57	65	75	83
South Bend, Ind.*	..	45	50	55	60	65	70
Newton, Mass.	39	55	59	73	85	94	102

*The tests given in these cases were not the Holmes tests, but were so like them that the results are comparable.

TABLE XIII

HOLMES PENMANSHIP TEST

Average Speed in 1 and 4 Minute Tests

Grade	2	3	4	5	6	7	8
Cape Giradeau	38	66	75	100	112	111	127
	5	3	7	5	7	9	15
	20	*8*	*13*	*15*	*13*	*22*	*16*
Columbia	35	32	51	87	89	92	
	4	12	10	11	13	12	
	11	*10*	*17*	*11*	*11*	*14*	*0*
Kirksville	31	39	59	69	83	71	87
	12	9	14	9	7	12	28
	11	*15*	*15*	*12*	*21*	*17*	*6*
Maryville	39	54	51	55	83	89	101
	4	12	9	9	16	18	9
	6	*10*	*10*	*16*	*8*	*12*	*10*
Springfield	44	69	75	97	110	96	106
	4	11	9	12	8	9	9
	14	*16*	*14*	*10*	*15*	*17*	*14*
Warrensburg	45	44	60	79	82	92	92
	9	7	10	18	19	8	10
	13	*21*	*15*	*20*	*21*	*41*	*31*
All Schools	40	49	62	80	92	92	102
	6	15	12	19	16	13	16
	75	*80*	*84*	*84*	*89*	*123*	*77*
Newton, Mass.	39	55	59	73	85	94	102
Brookline, Mass.	..	..	..	76	87	90	98
				10	9	13	13
				324	*330*	*309*	*273*

TABLE XIV

HOLMES PENMANSHIP TEST

Quality of Handwriting in Speed Test

Grade	2	3	4	5	6	7	8
Cape Giradeau	30	33	35	30	30	40	38
	3	5	3	3	3	5	3
	22	*6*	*13*	*15*	*13*	*22*	*16*
Columbia	48	45	45	43	42	46	
	5	9	3	4	4	6	
	11	*10*	*18*	*11*	*11*	*14*	*0*
Kirksville	33	33	35	35	35	36	33
	5	7	7	5	8	9	5
	19	*15*	*16*	*12*	*22*	*18*	*9*
Maryville	32	40	38	40	43	38	45
	3	7	8	10	7	7	10
	6	*10*	*10*	*16*	*8*	*12*	*10*
Springfield	30	25	26	33	36	33	38
	6	5	3	5	4	7	8
	16	*16*	*14*	*10*	*15*	*17*	*14*
Warrensburg	38	40	42	45	45	48	45
	7	5	6	5	7	7	5
	15	*21*	*15*	*20*	*21*	*41*	*31*
All Schools	33	35	35	38	38	43	42
	5	7	7	7	7	9	7
	89	*80*	*76*	*84*	*90*	*124*	*80*
Cleveland	..	..	..	45	48	50	55
Starch Standards	27	33	37	43	47	53	57
South Bend, Ind.	..	40	40	50	50	60	60
Newton, Mass	54	50	45	48	51	50	53

TABLE XV

HOLMES PENMANSHIP TEST

Quality of Handwriting in Dictation Test

Grade	2	3	4	5	6	7	8
Cape Giradeau	35	38	47	53	48	53	60
	7	5	7	5	2	7	7
	22	*8*	*13*	*15*	*13*	*22*	*16*
Columbia	45	47	42	43	47	55	
	3	7	5	3	6	10	
	12	*10*	*18*	*11*	*11*	*14*	
Kirksville	35	30	43	38	40	40	45
	5	3	5	5	5	7	5
	18	*15*	*16*	*12*	*24*	*18*	*9*
Maryville	27	38	38	38	48	40	58
	2	5	3	5	6	5	13
	6	*10*	*10*	*16*	*8*	*12*	*10*
Springfield	35	30	35	37	43	43	55
	5	3	3	5	6	7	12
	16	*16*	*14*	*10*	*15*	*17*	*14*
Warrensburg	37	40	45	53	50	55	50
	3	5	3	7	8	7	5
	15	*21*	*14*	*20*	*21*	*39*	*31*
All Schools	35	35	40	43	45	50	52
	7	7	5	8	7	10	8
	89	*80*	*75*	*84*	*92*	*122*	*90*

TABLE XVI

HOLMES PENMANSHIP TEST

Quality of Handwriting in Reproduction Test

Grade	2	3	4	5	6	7	8
Cape Giradeau	30	35	40	40	45	43	45
	3	3	7	3	7	5	7
	22	*8*	*12*	*15*	*12*	*21*	*16*
Columbia	52	38	48	47	45	48	
	3	6	5	7	3	5	
	12	*10*	*17*	*11*	*12*	*13*	*0*
Kirksville	33	38	38	36	40	40	53
	5	7	5	6	5	10	5
	12	*14*	*16*	*10*	*23*	*14*	*5*
Maryville	38	38	38	43	50	35	57
	5	5	10	7	7	10	15
	6	*10*	*10*	*16*	*8*	*12*	*10*
Springfield	37	28	38	40	47	42	43
	5	5	5	7	5	4	10
	14	*15*	*14*	*12*	*16*	*17*	*14*
Warrensburg	35	38	40	43	48	50	45
	5	5	3	7	8	5	5
	16	*21*	*15*	*20*	*21*	*41*	*32*
All Schools	35	35	40	42	45	45	45
	5	5	5	6	7	7	7
	82	*78*	*74*	*84*	*92*	*118*	*77*

TABLE XVII

HOMLES PENMANSHIP TEST

Average Quality of Handwriting

Grade	2	3	4	5	6	7	8
Cape Giradeau	33	36	43	39	42	47	47
	5	3	5	4	4	5	4
	22	*8*	*12*	*15*	*13*	*22*	*16*
Columbia	50	42	46	43	44	49	
	4	6	5	3	6	4	
	12	*10*	*18*	*11*	*11*	*14*	*0*
Kirksville	35	32	39	36	38	35	42
	3	6	4	5	4	5	7
	18	*15*	*16*	*12*	*22*	*18*	*9*
Maryville	32	41	38	43	45	37	54
	4	5	6	7	9	7	9
	6	*10*	*10*	*16*	*8*	*12*	*10*
Springfield	34	29	33	36	43	40	44
	4	3	4	6	5	4	9
	16	*16*	*14*	*10*	*15*	*17*	*14*
Warrensburg	35	39	43	48	47	51	45
	1	4	4	8	9	6	4
	15	*21*	*15*	*20*	*21*	*41*	*31*
All Schools	35	36	39	41	42	45	47
	6	6	6	5	5	8	6
	89	*80*	*75*	*84*	*90*	*124*	*80*
Newton, Mass.	54	50	45	48	51	50	53
Starch Standards	27	33	37	43	47	53	57
South Bend, Ind	..	40	40	50	50	60	60
Cleveland	..	..	..	45	48	50	55

TABLE XVIII

COMPARISON OF QUALITY OF PENMANSHIP ON DIFFERENT TESTS

Grade	2	3	4	5	6	7	8
Grade Medians in							
1. Speed Test							
	33	35	35	38	38	43	42
	5	7	7	7	7	9	7
	89	*80*	*76*	*84*	*90*	*124*	*80*
2. Dictation Test							
	35	35	40	43	45	50	52
	7	7	5	8	7	10	8
	89	*80*	*75*	*84*	*92*	*122*	*80*
3. Reproduction							
	35	35	40	42	45	45	45
	5	5	5	6	7	7	7
	82	*78*	*74*	*84*	*92*	*118*	*77*
4. Average of Three Tests							
	35	36	39	41	42	45	47
	5	6	6	5	5	8	6
	89	*80*	*75*	*84*	*90*	*124*	*80*

TABLE XIX

KANSAS SILENT READING TEST

Grade	3	4	5	6	7	8
Cape Giradeau	18.1	16.3	23.8	10.6	30.0	23.7
	9.5	4.6	4.8	2.4	4.2	7.5
	8	*13*	*14*	*13*	*21*	*15*
Columbia	6.8	12.4	16.3	15.4	16.3	
	3.0	5.2	4.4	4.6	5.1	
	10	*19*	*11*	*12*	*14*	*0*
Kirksville	5.2	9.8	12.8	16.8	19.3	9.5
	2.1	3.0	5.8	3.3	7.1	1.0
	14	*13*	*11*	*18*	*16*	*3*
Maryville	7.2	8.5	14.6	11.7	15.8	14.1
	3.2	4.5	7.3	2.1	4.1	0.7
	10	*9*	*16*	*8*	*12*	*8*
Springfield	10.5	13.4	18.3	13.5	17.3	18.4
	4.0	1.5	3.3	3.9	6.3	3.9
	15	*15*	*12*	*16*	*17*	*14*
Warrensburg	10.3	10.3	12.5	12.9	20.1	18.7
	6.0	3.9	3.8	2.5	4.3	5.8
	20	*15*	*19*	*21*	*41*	*28*
All Schools	8.5	12.4	16.3	13.2	20.0	18.4
	4.3	4.9	5.3	3.8	6.2	5.3
	77	*84*	*83*	*88*	*121*	*68*
Scores in 19 cities	6.0	9.9	13.7	13.4	16.5	18.8
	1207	*1473*	*1535*	*1264*	*1338*	*912*

TABLE XX

RATE OF SILENT READING

Judd Material*

Grade	2	3	4	5	6	7	8
Cape Giradeau	79	250	127	184	193	298	309
	3	78	28	62	45	32	12
	22	*8*	*12*	*15*	*13*	*21*	*16*
Columbia	133	189	156	184	272	236	
	60	51	57	32	48	49	
	12	*10*	*19*	*11*	*14*	*12*	*0*
Kirksville	102	144	154	213	210	198	261
	52	70	57	12	35	36	60
	14	*13*	*16*	*10*	*22*	*16*	*4*
Maryville	107	149	188	169	237	212	219
	43	60	60	54	57	47	54
	6	*10*	*10*	*16*	*8*	*12*	*10*
Springfield	122	172	201	251	241	184	200
	43	50	41	28	44	30	45
	16	*15*	*14*	*12*	*16*	*16*	*13*
Warrensburg	99	148	176	216	192	212	209
	51	65	80	51	58	50	45
	15	*21*	*15*	*20*	*20*	*41*	*26*
All Schools	104	169	166	202	216	224	235
	50	. 66	57	51	53	53	56
	85	*77*	*86*	*84*	*95*	*118*	*69*
Cleveland†	114	159	165	165	179	168	187
Standard Rates‡	90	138	132	154	167	161	172

The above scores are the average rates of silent reading expressed in number of words read per minute. The figures in the second line are the average deviations.

---

*This material was not used with the separate individuals as in the Cleveland Survey, but the test was given like the Holmes Test to the whole class. The results are, therefore, not exactly comparable.

†These figures estimated from diagram p. 152 of Measuring the Work of the Public Schools. Figures are averages obtained from scores of 1831 Cleveland pupils.

‡These scores are given on the Score Sheet for Reading published by the School of Education at the University of Chicago. They are also on the diagram referred to on page (21). They include scores of 2654 pupils in 13 cities.

TABLE XXI

HOLMES READING TEST

Speed of Silent Reading

Grade	2	3	4	5	6	7	8
Cape Giradeau	147	384	237	222	234	463	444
	60	180	30	42	57	100	81
	22	*8*	*12*	*15*	*13*	*20*	*15*
Columbia	186	147	219	309	327	324	
	96	33	54	33	97	39	
	12	*10*	*19*	*11*	*12*	*13*	*0*
Kirksville	126	207	180	321	264	276	366
	21	90	60	74	57	68	21
	19	*13*	*16*	*10*	*23*	*16*	*3*
Maryville	180	180	405	363	363	291	309
	30	66	153	96	81	72	75
	6	*10*	*10*	*15*	*8*	*12*	*10*
Springfield	138	237	188	363	276	237	276
	24	123	44	84	45	51	86
	15	*15*	*14*	*12*	*16*	*17*	*14*
Warrensburg	141	180	147	299	249	309	384
	63	33	24	67	42	72	123
	15	*21*	*15*	*20*	*20*	*41*	*28*
All Schools	147	198	198	309	276	318	363
	39	84	51	84	60	81	86
	89	*77*	*86*	*83*	*92*	*119*	*70*
Brookline, Mass				249	283	309	294
				69	80	63	69
				326	*339*	*313*	*273*
Fall River, Mass.					270	309	
					75	72	
					352	*312*	

TABLE XXII

HOLMES READING TEST

Reproduction of Story Read by Pupils

Grade	2	3	4	5	6	7	8
Cape Giradeau	14	66	60	63	63	69	66
	9	17	11	6	11	8	11
	22	*8*	*12*	*15*	*12*	*21*	*16*
Columbia	29	37	29	57	49	57	
	12	9	14	9	6	11	
	12	*10*	*17*	*11*	*12*	*13*	*0*
Kirksville	11	43	40	51	60	63	
	8	20	17	9	11	6	
	10	*12*	*16*	*10*	*23*	*14*	*0*
Maryville	41	29	43	51	57	57	63
	17	16	11	5	12	9	11
	8	*10*	*10*	*15*	*8*	*12*	*10*
Springfield	26	49	54	49	57	54	54
	9	17	6	6	9	11	12
	14	*15*	*14*	*12*	*16*	*17*	*14*
Warrensburg	14	29	43	51	51	57	57
	8	12	9	11	12	6	11
	14	*21*	*15*	*20*	*20*	*41*	*28*
All Schools	20	37	43	51	54	60	60
	11	14	14	8	9	11	11
	78	*76*	*84*	*83*	*85*	*118*	*68*
Brookline, Mass.				57	57	57	60
				9	9	9	9
				317	*329*	*312*	*269*
Fall River, Mass.					54	57	..
					9	9	
					352	*312*	

TABLE XXIII

HOLMES READING TEST

Questions on Story Read by Pupil

Grade	2	3	4	5	6	7	8
Cape Giradeau	39	71	77	73	77	85	77
	14	9	7	7	5	4	11
	12	*8*	*12*	*15*	*12*	*21*	*16*
Columbia	53	53	44	68	70	73	
	4	4	11	5	7	5	
	12	*10*	*17*	*11*	*12*	*13*	*0*
Kirksville	19	39	49	64	69	68	83
	19	15	8	10	8	14	6
	19	*12*	*15*	*10*	*22*	*17*	*3*
Maryville	44	49	54	69	72	69	72
	6	11	11	12	7	14	15
	6	*10*	*10*	*14*	*8*	*12*	*10*
Springfield	38	59	68	77	78	77	82
	9	5	10	14	5	8	12
	16	*15*	*14*	*12*	*16*	*17*	*14*
Warrensburg	19	44	44	72	68	68	72
	15	5	9	13	15	9	8
	15	*21*	*15*	*20*	*20*	*41*	*27*
All Schools	37	50	52	69	72	72	75
	16	12	15	8	7	11	10
	90	*76*	*83*	*82*	*90*	*121*	*70*
Brookline, Mass.				62	64	68	73
				11	11	9	9
				325	*340*	*313*	*274*
Fall River, Mass.					66	69	
					11	9	
					352	*312*	

TABLE XXIV

HOLMES READING TEST

Reproduction of Story Read to Pupils

Grade	2	3	4	5	6	7	8
Cape Giradeau	17	46	51	54	57	74	69
	11	15	5	8	6	9	12
	22	*8*	*11*	*14*	*13*	*22*	*14*
Columbia	26	37	46	57	71	60	..
	6	12	11	11	8	9	
	11	*9*	*19*	*11*	*12*	*14*	*0*
Kirksville	..	14	43	34	51	60	51
		8	12	0	5	14	8
		16	*17*	*13*	*25*	*19*	*8*
Maryville	40	43	46	46	54	60	54
	11	6	11	8	6	6	11
	6	*10*	*9*	*16*	*8*	*12*	*8*
Springfield	29	49	51	51	60	60	66
	9	8	2	3	9	6	8
	16	*12*	*13*	*12*	*16*	*17*	*14*
Warrensburg	26	34	40	54	54	57	63
	20	12	9	11	9	11	11
	15	*21*	*15*	*20*	*21*	*41*	*30*
All Schools	26	40	49	51	57	60	63
	11	11	8	9	7	9	9
	70	*68*	*73*	*78*	*78*	*112*	*70*

The above scores are in median per cents and the figures in the second line are the median deviations.

No standards for comparison are available, as this test has not been given previously. It was given to determine whether children can remember better what they hear or what they read. See Table XXXV for comparisons.

TABLE XXV

HOLMES READING TEST

Questions on Story Read to Pupils

Grade	2	3	4	5	6	7	8
Cape Giradeau	31	66	75	74	71	85	86
	14	8	10	11	9	5	6
	21	*8*	*12*	*14*	*13*	*22*	*14*
Columbia	48	62	60	70	83	76	
	10	6	10	5	4	4	
	11	*9*	*19*	*10*	*12*	*14*	*0*
Kirksville	..	37	62	67	68	67	78
		17	20	2	12	8	3
	0	*16*	*17*	*13*	*25*	*19*	*4*
Maryville	53	47	64	66	74	79	79
	22	16	23	5	8	8	9
	6	*10*	*9*	*16*	*8*	*12*	*8*
Springfield	48	65	69	85	85	86	88
	9	5	5	5	4	6	8
	16	*12*	*13*	*12*	*16*	*17*	*14*
Warrensburg	23	48	49	67	75	77	76
	15	13	17	7	7	7	7
	15	*21*	*15*	*20*	*21*	*41*	*30*
All Schools	37	54	66	69	75	79	80
	14	13	11	7	8	7	7
	69	*68*	*74*	*77*	*78*	*112*	*70*

The above scores are in median per cents and the figures in the second line are the median deviations.

No standards for comparison are available, as this test has not been given previously. It was given to determine whether children can remember better what they hear or what they read. See Table XXXVI for comparisons.

TABLE XXVI

HOLMES READING TEST

Comparison of Reproductions of Stories Read by and Read to Pupils

Grade	2		3		4		5		6		7		8	
	Read by	Read to	Read by	Read to	Read by	Read to	Read by	Read to	Read by	Read to	Read by	Read to	Read by	Read to
Cape Giradeau	14	17	66	46	60	51	63	54	63	57	69	74	66	69
	9	11	17	15	11	5	6	8	11	6	8	9	11	12
Columbia	29	27	37	37	29	46	57	57	40	71	57	69	no pupils	
	12	6	9	12	14	11	9	11	6	8	11	9		
Kirksville	11	no	43	14	40	43	51	34	60	51	63	60	no pupils	
	8	pupils	20	12	17	12	9	0	11	5	6	14		
Maryville	41	40	29	43	43	46	51	46	57	54	59	60	63	54
	17	11	16	6	11	11	5	8	8	6	7	6	11	11
Springfield	26	29	49	49	54	51	49	51	57	60	54	60	54	66
	9	9	17	8	6	2	6	3	9	9	11	6	12	8
Warrensburg	14	26	29	34	43	40	51	54	51	54	57	57	57	63
	8	20	12	12	9	9	11	11	12	9	6	11	11	11
All Schools	20	26	37	40	43	49	51	51	54	57	60	60	60	60
	11	11	14	11	14	8	8	9	9	7	11	9	11	9

Scores are median percentages with median deviations beneath in second line

TABLE XXVII

HOLMES READING TEST

Comparison of Answers to Questions on Stories Read by and Read to Pupils

Grade	2		3		4		5		6		7		8	
	Read by	Read to	Read by	Read to	Read by	Read to	Read by	Read to	Read by	Read to	Read by	Read to	Read by	Read to
Cape Giradeau	39	31	71	66	77	75	73	74	77	71	85	85	77	86
	17	14	9	8	7	10	7	11	5	9	4	5	11	6
Columbia	53	48	53	62	44	60	68	70	70	83	73	76	..	..
	4	10	4	6	11	10	5	5	7	4	5	4	..	..
Kirksville	19	..	39	37	49	62	64	67	69	68	68	67	83	80
	19	..	15	17	8	20	10	2	8	12	14	8	6	5
Maryville	44	53	49	47	54	64	69	66	72	74	69	79	72	79
	6	22	11	16	11	23	12	5	7	8	14	8	15	9
Springfield	38	48	59	65	68	69	77	85	78	85	78	86	82	88
	9	9	5	5	10	5	14	5	5	4	8	6	12	8
Warrensburg	19	23	44	48	44	49	72	67	68	75	68	77	72	76
	15	15	5	13	9	17	13	7	15	7	9	7	8	7
All Schools	37	37	50	54	52	66	69	69	72	75	72	79	75	80
	16	14	12	13	15	11	8	7	7	8	11	7	10	7

Scores are median percentages with median deviations beneath in second line

TABLE XXVIII
COMPOSITIONS

Grade	6	7	8
Cape Giradeau	..	74	75
		9	8
		21	*15*
Columbia	..	70	..
		9	
		14	
Kirksville	..	64	72
		5	8
		16	*4*
Maryville	..	67	69
		7	9
		12	*9*
Springfield	..	75	75
		3	8
		17	*14*
Warrensburg	..	57	65
		6	9
		35	*29*
All Schools	..	67	70
		8	10
		115	*70*
Bloomington, Ind.*	..	61	67
		8	8
		139	*129*
Port Townsend, Wash.	..	53	58
Newton, Mass.	..	..	75
			25
Brookline, Mass.	61	..	70
	6		7
	330		*278*

*These medians and variations were found by combining the scores of the A and B divisions of the grades in Bloomington. See Report of Second Indiana Conference on Educational Measurements, p. 117ff.

EXPLANATION OF ABBREVIATIONS USED IN PROGRESS TABLES
on the Following Pages

(a) Courtis Addition
(b) Courtis Subtraction
(c) Courtis Multiplication
(d) Courtis Division
(e) Stone Arithmetic
(f) Ayres' Spelling
(g) Speed of Penmanship
(h) Quality of Penmanship
(i) Kansas Silent Reading
(j) Speed of Silent Reading
(k) Quality of Reproduction
(l) Answers to Questions on Story Read

TABLE XXIX

PROGRESS IN WHOLE GROUP

Grade	2	3	4	5	6	7	8
(a)*	..	3–0	4–1	6–2	6–2	7–4	8–4
		1–0	1–1	1–2	1–2	1–2	1–2
(b)*	..	2–0	4–1	6–4	6–3	8–6	10–6
		1–0	1–1	1–2	2–2	2–2	2–2
(c)*	..	"O"	2–0	5–3	5–2	6–4	9–4
			2–0	1–2	2–2	2–2	2–2
(d)*	..	"O"	1–0	4–1	4–2	5–3	8–6
			1–0	1–1	2–2	2–2	3–2
(e)	..	1.0	1.6	3.0	4.0	6.4	7.0
		1.0	1.4	2.0	1.6	1.6	2.0
(f)	81	74	64	78	74	73	77
	17	15	23	18	18	19	16
(g)	40	49	62	80	92	92	102
	6	15	12	19	16	13	16
(h)	35	36	39	41	42	45	47
	6	6	6	5	5	8	6
(i)	..	8.5	12.4	16.3	13.2	20.0	18.4
		4.3	4.9	5.3	3.8	6.2	5.3
(j)	147	198	198	309	276	318	368
	39	84	51	84	60	81	86
(k)	20	37	43	51	54	60	60
	11	14	14	8	8	11	11
(l)	37	50	52	69	72	72	75
	16	12	15	8	7	11	10

*The figure at the left of the dash shows the median number of examples attempted, the figure after the dash shows the median number of examples right.

TABLE XXX

PROGRESS IN CAPE GIRADEAU

	Grade 2	3	4	5	6	7	8	Score in
(a)	..	3–0	4–0	6–3	8–3	8–5	9–7	Number of Examples
		1–0	1–2	1–1	2–3	2–2	2–2	
		9	*12*	*15*	*13*	*21*	*15*	
(b)	..	3–0	4–2	8–5	8–5	8–7	9–6	
		1–0	1–1	1–1	2–3	2–2	2–2	
(c)	..	1–0	3–0	5–3	8–4	7–4	10–6	
		1–0	2–0	1–1	1–2	2–2	1–2	
(d)	..	"0"	1–0	4–1	5–3	5–4	8–7	
			1–0	1–1	3–2	2–2	4–4	
(e)	..	1.0	2.0	3.0	5.0	7.1	6.9	Assigned values
		1.0	1.0	1.0	2.0	1.0	3.1	
		8	*13*	*15*	*13*	*22*	*16*	
(f)	76	84	82	85	80	83	82	Percent.
	14	8	13	12	10	13	12	
	21	*8*	*11*	*14*	*13*	*22*	*14*	
(g)	38	66	75	100	112	111	127	Letters per minute
	5	3	7	5	7	9	15	
	20	*8*	*13*	*15*	*13*	*22*	*16*	
(h)	33	36	43	39	42	47	47	Grade on Ayres' Scale
	5	3	5	4	4	5	4	
	22	*8*	*12*	*15*	*13*	*22*	*16*	
(i)	..	18.1	16.3	23.8	10.6	30.0	23.7	Assigned values
		9.5	4.6	4.8	2.4	4.2	7.5	
		8	*13*	*14*	*13*	*21*	*15*	
(j)	147	384	237	222	234	463	444	Words per minute
	60	180	30	42	57	100	81	
	22	*8*	*12*	*15*	*13*	*20*	*15*	
(k)	14	66	60	63	63	69	66	Percent.
	9	17	11	6	11	8	11	
	12	*8*	*12*	*15*	*12*	*21*	*16*	
(l)	39	71	77	73	77	85	77	Percent.
	14	9	7	7	5	4	11	
	22	*8*	*12*	*15*	*12*	*21*	*16*	

TABLE XXXI

PROGRESS IN COLUMBIA

Grade	2	3	4	5	6	7	8
(a)	..	2–0	3–0	5–2	7–2	6–4	
		1–0	1–0	1–1	2–1	2–1	
		10	*18*	*11*	*12*	*14*	
(b)	..	1–0	1–0	4–2	6–3	7–6	
		1–0	1–0	1–1	2–2	1–1	
(c)	..	"0"	1–0	1–0	3–0	4–2	
			0–0	0–0	1–0	1–2	
(d)	..	"0"	"0"	1–0	2–1	5–2	
				"0"–"0"	1–1	2–1	
(e)	..	1.0	2.5	4.6	4.6	5.0	
		0.4	2.1	1.0	1.6	1.7	
		10	*18*	*11*	*14*	*11*	
(f)	..	60	57	66	67	76	
		17	25	21	25	11	
		10	*19*	*11*	*12*	*14*	
(g)	35	32	51	87	89	92	
	4	12	10	11	13	12	
	11	*15*	*15*	*12*	*21*	*17*	
(h)	50	42	46	43	44	49	
	4	6	5	3	6	4	
	12	*10*	*18*	*11*	*11*	*14*	
(i)	..	6.8	12.4	16.3	15.4	16.3	
		3.0	5.2	4.4	4.6	5.1	
		10	*19*	*11*	*12*	*14*	
(j)	186	147	219	309	327	324	
	96	33	54	33	97	39	
	12	*10*	*19*	*11*	*12*	*13*	
(k)	29	37	29	57	49	57	
	12	9	14	9	6	11	
	12	*10*	*17*	*11*	*12*	*13*	
(l)	53	53	44	68	70	73	
	4	4	11	5	7	5	
	12	*10*	*17*	*11*	*12*	*13*	

No eighth grade in this school

TABLE XXXII

PROGRESS IN KIRKSVILLE

Grade	2	3	4	5	6	7	8
(a)	..	4–1	4–1	5–1	6–2	6–2	10–1
		1–0	1–0	1–1	2–1	2–2	4–1
		16	*17*	*12*	*22*	*18*	*8*
(b)	..	1–0	4–0	6–2	5–1	6–3	8–3
		1–0	1–0	3–2	1–1	2–2	2–1
(c)	..	"0"	1–0	3–1	3–0	6–2	10–0
			1–0	2–1	1–0	1–1	3–0
(d)	..	"0"	1–0	2–0	3–1	4–2	4–1
			1–0	1–0	2–1	1–1	4–1
(e)	..	0	1.0	1.0	4.0	5.2	4.0
		0	1.0	1.0	2.0	1.2	1.0
		15	*15*	*13*	*22*	*18*	*9*
(f)	67	66	53	73	64	58	57
	22	20	29	20	19	24	21
	17	*14*	*16*	*10*	*24*	*18*	*8*
(g)	31	39	59	69	83	71	87
	12	9	14	9	7	12	28
	11	*15*	*15*	*12*	*21*	*17*	*6*
(h)	35	32	39	36	38	35	42
	3	6	4	5	4	5	7
	18	*15*	*16*	*12*	*22*	*18*	*9*
(i)	..	5.2	9.8	12.8	16.8	19.3	9.5
		2.1	3.0	5.8	3.3	7.1	1.0
		14	*13*	*11*	*18*	*16*	*3*
(j)	126	207	180	321	264	276	366
	21	90	60	74	57	68	21
	19	*13*	*16*	*10*	*23*	*16*	*3*
(k)	11	43	40	51	60	63	no
	8	20	17	9	11	6	pupils
	19	*12*	*15*	*10*	*22*	*17*	
(l)	19	39	49	64	69	68	83
	19	15	8	10	8	14	6
	19	*12*	*15*	*10*	*22*	*17*	*3*

TABLE XXXIII

PROGRESS IN MARYVILLE

Grade	2	3	4	5	6	7	8
(a)	..	5–1	6–4	6–4	7–5	8–3	8–5
		1–1	1–2	1–2	2–2	1–1	1–1
		9	*12*	*15*	*13*	*21*	*15*
(b)	..	2–0	5–4	6–4	9–5	9–6	10–7
		1–0	2–1	1–2	2–2	1–2	1–1
(c)	..	3–0	5–1	5–3	7–4	8–5	8–3
		1–0	1–2	1–1	2–1	1–2	2–2
(d)	..	"0"	3–1	4–2	5–3	6–3	6–4
			2–1	1–1	3–2	1–1	2–1
(e)	..	0	0.5	3.1	4.0	6.7	6.3
		0	0.5	2.0	1.2	0.7	1.3
		10	*10*	*16*	*8*	*12*	*10*
(f)	97	93	73	80	76	76	90
	2	8	17	18	17	20	12
	6	*10*	*9*	*16*	*8*	*12*	*10*
(g)	39	54	51	55	83	89	101
	4	12	9	9	16	18	9
	6	*10*	*10*	*16*	*8*	*12*	*10*
(h)	32	41	38	43	45	37	42
	4	5	6	7	9	7	7
	6	*10*	*10*	*16*	*8*	*12*	10
(i)	..	7.2	8.5	14.6	11.7	15.8	14.1
		3.2	4.5	7.3	2.1	4.1	0.7
		10	*9*	*16*	*8*	*12*	*8*
(j)	180	180	405	363	363	291	309
	30	66	153	96	81	72	75
	6	*10*	*10*	*15*	*8*	*12*	*10*
(k)	41	29	43	51	57	57	63
	17	16	11	5	12	9	11
	8	*10*	*10*	*15*	*8*	*12*	*10*
(l)	44	49	54	69	72	69	72
	6	11	11	12	7	14	15
	6	*10*	*10*	*14*	*8*	*12*	*10*

TABLE XXXIV

PROGRESS IN SPRINGFIELD

Grade	2	3	4	5	6	7	8
(a)	..	2–0	4–1	6–3	6–4	6–4	8–3
		1–0	1–1	2–2	2–2	1–1	2–2
		14	*15*	*12*	*16*	*16*	*14*
(b)	..	4–1	4–1	7–5	8–6	8–5	10–6
		1–1	1–1	0–2	2–2	2–2	2–2
(c)	..	"0"	3–1	5–4	8–5	6–3	8–4
			1–1	1–1	2–1	1–1	3–2
(d)	..	"0"	2–0	3–2	6–4	5–4	9–7
			0–0	1–1	1–2	1–1	1–2
(e)	..	1.0	2.0	3.6	4.6	6.8	6.8
		0.6	1.0	3.6	1.0	1.2	1.8
		16	*15*	*10*	*15*	*17*	*14*
(f)	95	85	72	89	81	79	76
	5	8	15	8	15	17	19
	16	*16*	*14*	*12*	*16*	*17*	*14*
(g)	44	69	75	97	110	96	106
	4	11	9	12	8	9	9
	14	*16*	*14*	*10*	*15*	*17*	*14*
(h)	34	29	33	36	43	40	44
	4	3	4	6	5	4	9
	16	*16*	*14*	*10*	*15*	*17*	*14*
(i)	..	10.5	13.4	18.3	13.5	17.3	18.4
		4.0	1.5	3.3	3.9	6.3	3.9
		15	*15*	*12*	*16*	*17*	*14*
(j)	138	237	188	363	276	237	276
	24	123	44	84	45	51	86
	15	*15*	*14*	*12*	*16*	*17*	*14*
(k)	26	49	54	49	57	54	54
	9	17	6	6	9	11	12
	14	*15*	*14*	*12*	*16*	*17*	*14*
(l)	38	59	68	77	78	78	82
	9	5	10	14	5	8	12
	15	*15*	*14*	*12*	*16*	*17*	*14*

TABLE XXXV

PROGRESS IN WARRENSBURG

Grade	2	3	4	5	6	7	8
(a)	..	3–0	5–1	6–3	6–2	8–4	8–4
		1–0	1–1	1–2	1–2	1–2	1–2
		20	*15*	*20*	*21*	*41*	*30*
(b)	..	3–0	4–2	7–4	5–3	9–7	10–7
		1–0	1–1	1–1	1–1	2–2	2–2
(c)	..	"0"	3–0	6–3	5–2	7–4	9–5
			1–0	1–2	2–2	1–2	2–2
(d)	..	"0"	2–0	5–3	4–2	6–4	7–6
			1–0	2–2	1–1	2–2	2–2
(e)	..	0	1.0	2.0	3.6	5.6	7.8
		0	1.0	1.0	1.8	2.0	1.4
		20	*15*	*20*	*21*	*41*	*32*
(f)	82	65	58	76	79	71	77
	15	17	22	20	17	20	19
	15	*21*	*16*	*20*	*21*	*41*	*30*
(g)	45	44	60	79	82	92	102
	9	7	10	18	19	8	16
	13	*21*	*15*	*20*	*21*	*41*	*31*
(h)	35	39	43	48	47	51	45
	1	4	4	8	9	6	4
	15	*21*	*15*	*20*	*21*	*41*	*31*
(i)	..	10.3	10.3	12.5	12.9	20.1	18.7
		6.0	3.9	3.8	2.5	4.3	5.8
		20	*15*	*19*	*21*	*41*	*28*
(j)	141	180	147	299	249	309	384
	63	33	24	67	42	72	123
	15	*21*	*15*	*20*	*20*	*41*	*28*
(k)	14	29	43	51	51	57	57
	8	12	9	11	12	6	11
	14	*21*	*15*	*20*	*20*	*41*	*28*
(l)	19	44	44	72	68	68	72
	15	5	9	13	15	9	8
	15	*21*	*15*	*20*	*20*	*41*	*29*

HARVARD STUDIES IN EDUCATION

Volume I: The Oberlehrer. A Study of the Social and Professional Evolution of the German Schoolmaster. By WILLIAM SETCHEL LEARNED. $1.25.

Volume II: The Appointment of Teachers in Cities. A Descriptive Critical and Constructive Study. By FRANK WASHINGTON BALLOU. $1.50.

Volume III: The Teaching of Economics in Harvard University. A Report Presented by the Division of Education at the Request of the Department of Economics. $2.00.

HARVARD BULLETINS IN EDUCATION

I. The School System as an Educational Laboratory. By WM. S. LEARNED. 25 cents.

II. Scales for the Measurement of English Composition. By FRANK W. BALLOU. 50 cents.

III. Bridging the Gap: The Transfer Class. By F. W. WRIGHT. 30 cents.

IV. A Selected Critical Bibliography of Vocational Guidance. By JOHN M. BREWER and ROY W. KELLY. 50 cents.

V. A Descriptive Bibliography of Measurement in Elementary Subjects. By HENRY W. HOLMES and others. 50 cents.

VI. Business Practice in Elementary Schools. By ROY DAVIS. 50 cents.

VII. Sight-Saving Classes in the Public Schools. By R. B. IRWIN. 35 cents.

Published by

THE GRADUATE SCHOOL OF EDUCATION
HARVARD UNIVERSITY, CAMBRIDGE 38, MASS.

HARVARD MONOGRAPHS IN EDUCATION

SERIES I

Studies in Educational Psychology and Educational Measurement

Manuscripts for Series I should be addressed to Professor Walter F. Dearborn, Psycho-Educational Clinic, Pulfrey House, Oxford Street, Cambridge 38, Mass.

Remittances should be made by check or money order to The Graduate School of Education, Harvard University, Cambridge 38, Mass.

Series I of the Harvard Monographs in Education has been established for publishing the results of statistical and experimental studies and of educational tests in the general fields of educational psychology and educational measurement.

The numbers are as follows:

1. A Comparison of the Intelligence and Training of School Children in a Massachusetts Town. 54 pp. E. A. Shaw and E. A. Lincoln. Postage prepaid, 60 cents.
2. The Marking System of the College Entrance Examination Board. 15 pp., 4 plates. L. Thomas Hopkins. Postage prepaid, 40 cents.
3. Standard Educational Tests in the Elementary Training Schools in Missouri. 90 pp. Walter F. Dearborn, Edward A. Lincoln, and Edwin A. Shaw. Postage prepaid, 75 cents.
4. A Series of Form Board and Performance Tests of Intelligence. Walter F. Dearborn, Edwin A. Shaw and Edward A. Lincoln. Postage prepaid, 75 cents.
5. The Determination of Anatomical Age in School Children and Its Relation to Mental Development. 60 pp., 19 plates. D. A. Prescott. Postage prepaid, 75 cents.

Whole No 4. Series 1, No 4.
HARVARD MONOGRAPHS IN EDUCATION

A SERIES OF FORM BOARD AND PERFORMANCE TESTS OF INTELLIGENCE

BY
WALTER F. DEARBORN
EDWIN A. SHAW
EDWARD A. LINCOLN

Series 1, No. 4

STUDIES IN EDUCATIONAL PSYCHOLOGY
AND
EDUCATIONAL MEASUREMENT

Edited by
WALTER F. DEARBORN

SEPTEMBER, 1923

Published by
THE GRADUATE SCHOOL OF EDUCATION
HARVARD UNIVERSITY, CAMBRIDGE, MASS.

FORM BOARD AND PERFORMANCE TESTS OF INTELLIGENCE

SECTION I—INTRODUCTION*

Form boards of the type described in this monograph were first devised as a means for the training of mentally deficient children. The ablest and most experienced teachers of the feebleminded from the days of Itard and Seguin, who were the originators of these methods, to the present time have relied on such means of training as are exemplified in the form boards. Montessori, noting the success of these methods with the deficient, believed that normal children might also profit by their use.

The psychologist regards learning as evidence of the functioning of intelligence. It was, therefore, a natural step for him to reverse the view of the educator and employ the devices for teaching as means of testing intelligence. This step appears to have been first taken by Dr. Naomi Norsworthy. She employed a "block test" as one of a series of tests in an investigation of the psychology of mentally deficient children.† These tests were given to children in the Massachusetts School for the Feebleminded at Waverley, and the design of the block test in question is doubtless to be traced to form boards constructed by Seguin then and still in use at this school.

Seguin had himself planned and partly constructed a series of form boards of increasing difficulty for school use. The tests herewithin described were first conceived as carrying out that intention, but primarily for the purpose of testing rather than that of training.‡

These tests have been in use for a number of years in the laboratory courses in educational psychology, and more recently in the work of the Psycho-Educational Clinic of Harvard University. They have been used chiefly to supplement the Binet tests in the making of intelligence examinations, and the intention has been ultimately to arrange

*The authors have co-operated in nearly every phase of this study, but there has been some division of labor: the introduction (Section 1) has been written by W. F. Dearborn; E. A. Shaw is chiefly responsible for the standardization of the tests, and E. A. Lincoln for the statistical work.

†The Psychology of Mentally Deficient Children, by Naomi Norsworthy, Ph.D. Archives of Psychology No. 1, 1906. The Science Press, N.Y., pp. 25, 26.

‡See Form Board and Construction Tests of Mental Ability, by W. F. Dearborn, J. E. Anderson and A. O. Christiansen. Journal of Educ. Psych., Vol. VII, 1916, pp. 445-458.

them as a scale which might take the place of the more linguistic tests in the examination of foreign speaking subjects and those suffering from defects of speech and hearing, and to supplement other tests used in the examination of normal and defective children. Because, however, of the large amount of time and effort involved in the standardizing and scaling of any individual test, this project has not been carried through to completion. Furthermore, it was found that a number of modifications were necessary in the directions for giving and scoring the tests, and it is only very recently that these revisions have been completed so that the standardization could go forward.

Although the work on this series of tests is still far from complete, it has seemed well to gather in this monograph the descriptions of all the construction or performance tests devised by the writer and his collaborators, together with the revised directions for giving and scoring, and the results which have been obtained so far in the process of standardization. A thorough-going discussion of the general problems of performance tests and scales,* such as their value as measures of intelligence, will not be attempted in this monograph, but a few comments on the methods and on the results thus far attained may well preface their presentation.

Sylvester in his discussion of the Seguin form board remarks: "It (the form board) appeals to the child's interests, affording him a short and fascinating task which calls for his best efforts, and it helps to free him from the fear and self consciousness which often interfere seriously in a mental examination. At the same time the tests give the examiner a good general view of the child's mentality and it usually indicates more or less clearly the nature of his defects."† Similar comments have frequently been made by mental examiners in the use of various tests of the performance or puzzle type. They feel as they observe the concrete working out of the test that they get an insight into some of the mental operations of the subject which they may miss in other forms of tests.‡. These impressions are often given considerable weight in the

---

*For an able discussion of some of these problems the reader is referred to "A Scale of Performance Tests," by Pintner and Patterson. Appleton & Co., N.Y. 1921.

†Sylvester, R. H., The Form Board Test, Psychological Monographs. Vol. XV, No. 4, p. 1.

‡For example, the following statements are made in a recent article in regard to Form Board 1C of this series. "The way in which children attack the problem and work at it indicates that they are thinking it out. The test gives the examiner an insight into the way the child sets about solving new problems." R. E. Leaming in the Psychological Clinic, Vol. XIV, No. 7, 1922.

examiner's "qualitative" rating of the subject, although he cannot give them "quantitative" expressions in terms, e.g., of a mental age scale, etc. The object of the standardization of such a test should be to give an objective measure of these, as the writer believes, undoubtedly well grounded, subjective impressions. This has been attempted in the present study by counting the number of correct and incorrect moves, and making note of the time taken to complete the task. Individuals can be differentiated on this basis particularly if small differences in speed are made much of. This is a very common method of standardization, but the writer is in doubt in regard to its validity and practical usefulness.

Large differences in time and moves are perhaps significant in the majority of cases. All too frequently, however, the count misses what the examiner prizes most in his observation. One child by a rapid fire of random moves with some good luck completes a test—with a score of many moves and short time—but leaving the impression that he has learned little. Another may make as many moves in a much longer time, but because of the character of his final moves, show that he has at last caught the point and really solved the puzzle. A third subject may make but a few moves, all of which are correct, but for fear of making a false move will unduly extend the time of the test. Some of these individual variations can doubtless be given an objective rating by taking account of the relation of time and moves in second and third trials. This has not been attempted in a sufficient number of cases to report, but will be tried in subsequent attempts at standardization. What has been done in the present instance is to record the results in time and moves for various chronological ages, and to classify individuals only on the basis of considerable differences in these respects. The clinical examiner must still make mental note of differences in method of attack, etc., which he regards as significant.*

---

*The following statement made in regard to Form Board 1C of this series is of interest in this connection. "One of the interesting things to anticipate is whether the individual's discrimination of form and his observation will be sufficiently good to enable him not to confuse the hexagon and the diamond in the Dearborn Form Board as is done in such a large number of cases. This seems to be a very difficult bit of discrimination. Another observation which is valuable in rating the individual is his method of attacking the problem; whether he uses trial and error, simply moving the blocks around from one position to another until they can all be fitted in somewhere or whether he studies the situation carefully and follows a plan of action which he formulates for the solution. The time rating alone gives no indication as to the method. It is possible, and is in fact often the case, that the trial and error method, which must be considered an inferior method, can solve the tests more quickly than a well-planned system for solution. The reaction time of the individual figures in this. An individual who moves rapidly can make several incorrect moves, correct them and find the correct move, while an individual of slow reaction time but precision of thought and movement will not yet have completed the same move." R. E. Leaming: art. cit. p. 209.

The results presented in the following pages show that the intelligence and learning required in these particular tests are possessed by most children by the age of nine or ten. The tests would appear, therefore, useful chiefly for children of ten years or under. In order for the present to make at least a nominal difference between the results of the tests and "intelligence" or "mental" age, the standards are stated as "performance ages." Since, however, much older children and even adults, who, by other criteria, would be considered at least average for their ages, do very poorly or even fail to do the tests within generous time limits, we may be testing here a skill which for lack of proper training or experience has not been acquired by these individuals. It may, therefore, frequently be significant to know that a child of, e.g., twelve years of age, who has done passably well in the academic subjects in school, and who is now looking for a job demanding some manual dexterity, is notably deficient in these tests. Whether or not this gap in his experience is one which can be quickly filled, is a question which must await further investigation, but, if it appears consistently in a series of such performance tests, this fact can hardly fail to have some present significance.*

On the basis of the present and possibly faulty methods of scoring, really essential differences cannot be established on the average after age seven or eight. On this account it may be better to disregard age and establish a percentile distribution of an unselected group of various ages. Such a procedure would be tacitly arguing that these tests (as now scored) either test only very simple mechanical constructive abilities, and that more involved tests are necessary for higher ages, or that this sort of skill is not especially well developed by ordinary school training. It is commonly argued that schools favor the development of an abstract and verbal type of intelligence. The extended use of these tests strongly incline one to the belief that such special training may result in the development of highly abstract intelligence in the handling of verbal and numerical symbols with but little development

*A boy of ten recently examined at the Psycho-Educational Clinic secured an intelligence quotient of 140 in the Stanford-Binet tests. His success in the more strictly verbal tests was notable. He came from a scholarly home, spent a good part of his time outside of school in reading, and was not skillful in baseball or other outdoor games, if, indeed, he played them at all. His performance in the form board tests was so disappointing, that the inquiry was directed to his parents as to whether the boy was not missing a sensori-motor training of a type which it might be difficult for him to make up for in later years, and which most boys acquire without any particular training. The parents were inclined to agree with the "diagnosis" and hoped that a recently installed work-bench in the cellar would give a start in the right direction.

of a more mechanical and in some cases, as Thorndike adds, social intelligence. These tests certainly do not differentiate between higher grades or stages in the development of mechanical intelligence, but they can, perhaps, indicate the lack of such development. No psychologist would attempt to draw hard and fast lines between abstract, mechanical and social intelligence, but varying experiences will often weight the final product and some classification of individual skills can at least be profitably made.

Such a classification which shows insight into important individual differences has recently been made by R. E. Leaming. It is based on the results of mental, scholastic and performance tests of 600 fifteen year old children. The results are considered "with reference to their value in vocational guidance," and according to this investigation disclose "four distinct types of children, in conformity to which every child tested may be definitely classified." Current notions in regard to the distribution and correlation of mental abilities would lead one to expect variation about a single type rather than a number of types. The considerations presented above would, however, also argue for less correlation between performance tests and the current intelligence tests than might usually be anticipated. The four types and the author's comment on them are given in the following quotation.

"(1) The superior children who do well in both the performance tests and the intellectual tests (i.e., the Hard Directions, Arithmetic, Reading and Binet tests).

(2) The children who do poorly in both the performance tests and in the intellectual tests.

(3) The children who do well in the intellectual tests, but make indifferent or even poor records in the performance tests, and

(4) The children who give excellent, sometimes even remarkable exhibitions in the performance tests and yet do very badly with the intellectual tests.

Children of the first type, the superior children of general competency, may be expected to adjust themselves in any environment and to succeed at almost anything that they may attempt.

The children of the second type are inferior children in both intellectual and mechanical competency. To give them vocational guidance it would be necessary to discover some special ability which an ordinary battery of psychological tests cannot be expected to discover, or else they must be recommended to secure training in efficiency at some occupation requiring only a low level of mentality or skill.

Children of the third type have a fairly high intellectual level, but the performance tests disclose either deficiency of intelligence in the sense of the ability to solve what for the individual is a new problem, or a deficiency in mechanical skill, the ability to work expeditiously with concrete material. Shall we advise a child of the third type to continue his school work and seek some intellectual field of activity? If the deficiency in performance is a deficiency of intelligence, surely not, but if it is a deficiency in mechanical skill a child of this group may be

expected to succeed at some profession requiring originality in intellectual operations. If, on the other hand, the deficiency in the performance tests is taken to indicate a lack of intelligence, then the child ought not to be advised to push on in school work. If he does, he will join the ranks of those who have been over-educated beyond their performance ability. They pass examinations and acquire knowledge, but are unable to use their acquired knowledge successfully in any occupation of high intellectual level. They disappoint their instructors and themselves. They join the ranks of failures in the intellectual occupations.

The children who conform to type 4 are not intellectual, but are either intelligent or skillful in mechanical operations. They should be made ready quickly, with a minimum of efficiency training, to take their place in the world in order that they may pit their intelligence and skill against the intelligence and skill of competitors in some productive field."*

The extent to which differences in training, experience and environment affect the "general intelligence" which our intelligence tests are supposed to measure has not been generally recognized by the mental tester. If we should accept Boring's definition that "intelligence is what the tests test,"** we are in need of different tests for different individuals. In the case of the influence of schooling on the standing in the Binet tests, Terman's assertion that "the lack of schooling does not prevent a subject from earning an average or superior score in the test"† will hold in such isolated cases as he presents in evidence of his statement, but probably not on the average, and will naturally be more often the case in the comparison of younger children, where the accumulative influence of schooling, or the lack of it, is less. Burt by the use of partial coefficients of correlation estimates that over one-half of the gross result is attributable to school attainment,‡ and that "linguistic ability and linguistic attainments exert upon the Binet-Simon tests a special and positive influence of their own.§ The preponderance given in current estimates of intelligence to word knowledge as compared to the kinaesthetic knowledge (which the form board tests may sample) will be referred to in the concluding paragraphs of this section.

Burt's statistical analysis of the effect of schooling on the test results finds ample support in the more recent study of Gordon¶ of phy-

---

*R. E. Leaming. op. cit. pp. 24-25. Philadelphia 1923.

**Intelligence as the Tests Test It. The New Republic. Vol. XXXV, No. 144 (June 6, 1923).

†Intelligence of School Children, p. 13.

‡Mental and Scholastic Tests, p. 183.

§Op. cit. p. 184.

¶Mental and Scholastic Tests among Retarded Children, by Hugh Gordon, Inspector of Schools; Board of Education, Educational Pamphlets, No. 44. London 1923.

sically defective, Gipsy, and Canal Boat children of little or no schooling. The number of cases studied by Gordon is not large, 84, 76, and 82 in the respective groups, but the accumulative evidence is convincing. Since the findings of Gordon have a direct bearing on the problems under discussion, and are not so well known as those of Burt, they may be summarized as follows: Occasional comments of the present writer are enclosed in parentheses.

The Stanford revision of the Binet tests was used for the testing of intelligence, and the results are expressed as "mental ratios," the ratio of the individual's mental age to his chronological age. Ballard's standardized tests in the speed of (discontinuous) reading, in the speed of adding and in the speed of subtraction were used to test the "scholastic" ability. A reading, adding, and subtraction "age" was determined for each child by comparing his accomplishment in the tests with the standards, which show what is the average accomplishment for each chronological age, and ratios calculated between his chronological age and his "test" age in each of the three subjects. The average of these three ratios is called the "educational ratio."

The average percentage of attendance of the physically defective children was 48% as compared with approximately 88% in the case of ordinary school children.

The average mental ratio and the average educational ratios were the same, + .86, and the correlation between the two ratios was + .78. Since the social environment and heredity of these children were of a character which would lead one to expect an average mental ratio, i.e., of 100, the lower mental ratio is probably due to the lack of schooling. Evidence is presented to show that the physical defects themselves, apart from the resulting loss of schooling, were not responsible for the result.

The correlation between the mental ratios and the attendance and between educational ratios and the attendance of these children was found to be the same, + .313 ± .066. *If the mental ratio is a measure of native ability, and the educational ratio of schooling, the effect of the lack of schooling should be greater in the latter than in the former ratio.*

The correlations between mental ratio and age and between educational ratio and age were negative — .42, ± .06 and — .288 ± .06 respectively, showing that the older the child the less his intelligence, and the less his scholastic ability in comparison with that which is normal for his age. (The fact that the relative decrease with increasing age is greater in the case of "intelligence" than in the case of scholastic ability will be considered a "fly" in the "ointment" of argument. It gains significance when taken in connection with findings in the case of children in "Backward Classes" of the Elementary Schools—who form a fourth group studied by Gordon in this report. The latter are "naturally dull children whose retardation appears to be in most cases due to a lack of native ability," and the more they "attend school the greater their scholastic ability, but the more they attend school the less their 'intelligence',"—as tested by the Binet tests. (pp. 73, 74.) In

explanation of these results it may be suggested that the limited intelligence of the backward children has been exhausted in the mechanics of reading and arithmetic, whereas in the case of the other groups studied, lack of schooling *plus* social isolation and the lack of the ordinary social and intellectual intercourse are responsible for the relatively greater deficiency in the tested intelligence especially in the case of the older children.)

More impressive are the comparisons in the case of the Canal Boat children. These children "in respect to health, cleanliness, morality, feeding, etc., are fully equal, if not superior, to town dwellers of a similar character. That they are not mentally defective, as is generally understood by that term, is shown by the life and wages of their parents, who in many cases have had no education and can neither read nor write. Their intellectual life, on the other hand, is of a most meagre description, owing to their lack of education and also owing to their social isolation." (pp. 35, 36.)

(In connection with the statement that these children are not mentally deficient, judging from the economic efficiency of their parents, it must be noted that the economic and social criteria even when applied in the case of city populations are less exacting than those commonly applied in the schools. In terms of the Binet tests an IQ of 50 or below [rather than of 70] would seem to correspond to the line of demarcation as established by economic or social standards. See article by the writer in School and Society, Vol XVII, No. 443, June 23, 1923, pp. 673-677.)

The schooling of these children is negligible; it probably averages between 20 and at most 40 half days as compared with 360 half day attendances of the child in the ordinary elementary school. In the school under review "it appears that the children attend about once a month for one to perhaps two and a half days. The maximum continuous attendance was five half days and such a number of attendances was very exceptional." The average attendance is estimated at about 4 to 5 per cent. as compared with 88% of children in ordinary elementary schools.

Only 36 out of 76 children tested were able to do all three of the scholastic tests; the mental ratios of these children average 71.5 and the educational ratios 70.4. The average mental ratios of the remaining 40 children was 67.9.

The correlation between the mental and educational ratios for the 36 children was 0.715, P. E. $\pm$ 0.054. For 57% of the children the difference between the mental ratio and educational ratio was only five points or less, and for 83% ten points or less." (p. 37.)

The rank correlation between the mental ratios and the chronological age of the children, the oldest being ranked first, was $-$ .755 $\pm$ .033. (In the case of ordinary school children the correlation between mental *ratio* and chronological age should approximate zero.)

(This negative correlation is particularly significant when taken in connection with an earlier statement (p. 20) that the children of six years of age or less in the group all had mental ratios between 90 and 100, that is, were of average intelligence, whereas with but one exception children over nine years of age scored 70 or below. Such a finding will be of interest in connection with recent discussions of racial differences

in intelligence, for example, the asserted relative brightness of the negro in the earlier years of schooling and his subsequent relative decrease in intelligence as measured by the tests in present use. The Canal Boat children are presumably of the same "stock" as the aggregate of English children with whom they are here compared through the test standards.)

The older children in the *same* family are, relative to their age, the less intelligent.

"There are twenty-two cases in which two or more children of the same family are attending this school—eleven families with two children, six with three children and five with four children."

"With one or two exceptions, an increase of age is found to be associated in the same family with a decrease in mental ratio." . . . "One significant fact is noticeable in many of the families, and it is that in the same family the mental ages are practically the same, although the chronological ages differ very considerably, e.g., 7-0, 6-3, 6-0; 5-6, 5-6; 6-9, 6-6; 6-0, and 6-3. This peculiarity seems to indicate that for children associating only with uneducated brothers and sisters there is a tendency to equalization of mental ages." (pp. 38, 39, 40.)

The average mental ratio of the youngest of four in a family (there were five such children with an average chronological age of 5.7) was + .87; the average mental ratio of the third children in family (11 cases) was approximately + .77, with an average chronological age of about 7.3; the average mental ratio of the second in family (22 cases) was about 73 with an average age of a little less than nine; and the average mental ratio of the oldest in family (22 cases) was + .60 with an average chronological age of 11.6. As Gordon concludes, these results make "it almost certain that the low average 'intelligence' of the children in this school is not due to heredity, seeing that the youngest children test more or less normally, i.e., are of average 'intelligence'."

"The tests found most difficult were those that might be expected to have been influenced by schooling, e.g., the vocabulary and drawing tests, and the making of sentences; those found the easiest were tests that are generally judged to be more independent of schooling, such as the absurdities, counting thirteen pennies, and repeating numbers backward." (p. 44.)

In general the reason why the younger children do better in the Binet tests is that the tests for the early years do not depend on schooling, and the failure of the older children is largely due to the fact that the tests for the higher ages do depend in large measure on schooling.

The possible usefulness of the tests described in the present monograph is indicated in a concluding paragraph. "Without the mental effort or mental exercises associated with schooling it would appear that there has been very little mental development on the intellectual side. How far there has been a similar lack of development among these children in connection with problems touching their own especial environment, it is difficult to say and, without tests especially devised and standardized for children in such surroundings, impossible to measure." (p. 45.)

Similar findings are reported in the tests on Gipsy children. The

percentage of school attendance was greater than in the case of Canal Boat children, about a fourth of them had less than 10% of attendance, and a half of them less than 30%, the average being about 35%. The average mental ratio of the 82 children tested was 74.5, their average age being nine years 9.5 months. The average mental and educational ratio of the 60 children who could do all of the scholastic tests were 75.4 and 77.4 respectively, and the correlation between these ratios + .784 ± .033. The correlations between mental ratio and age was (—) .430 ± .061. In the case of the 60 who could do the educational tests, the correlations were: mental ratio and attendance + .283 ± .08; educational and attendance + .289 ± .079; mental ratio and age — .566; educational ratio and age — .374. The findings in regard to the younger and older children of the same family parallel closely those in the case of Canal Boat children. The conclusions drawn from these various results are "that the mental tests used do not measure native ability apart from schooling except in the case of very young children, and, further, that for all practical school purposes simple standardized scholastic tests are of equal value." (p. 59.)

As Gordon adds, "the important question suggested by these results is whether there is any mental development apart from mental effort and such mental exercises as are generally associated with school life" and concludes that the answer "probably depends on the social environment of the children." We have found that ordinary school children do not improve much with age in the performance tests, but do in the Binet tests. Various possible explanations of this result have been suggested. Children with little or no schooling, according to Gordon's results, advance but little with age in the Binet tests. Would they show greater improvement with age in the performance tests? Gordon hazards the opinion that "the Gipsies might do better but not the Canal Boat children," and points to the value of an investigation "by means of specially devised environmental tests of possible mental development in other directions than those tested by the Binet tests."

Opinions will differ as to the relative rating in terms of intelligence of these various possible specializations in the subject matter of learning. One cannot, however, differentiate between grades of intelligence on the basis of the materials employed as, e.g., between the handling of "things" or "ideas," whether the latter are said to be concrete or abstract, or between the handling of things and the symbols for things, such as words and numbers. Bradley's discovery of the reason for the aberration of the light of the stars followed on the observation of a change in the position of the vane at the top of a boat's mast while he was sailing upon the river Thames.* The explanation of the change in

*"Our knowledge that light had a finite velocity followed on the invention of the telescope and the discovery of Jupiter's satellites: the news of their eclipses came late at times and these times were identified as those when Jupiter was un-

the position of the vane as due not to a change in the wind but to a change in the position of the boat was supplied by the sailors. Bradley applied the explanation to the apparent change in the direction of the light of a star. On the initial test, the intelligence of the sailors must be rated at least as equal to that of the scientist, for they held the explanation which he sought.

That the sailors reasoned of wind and boat and the scientist of light and stars would not seem a material difference in judging of the intelligences involved. But as a result of an added item of information the scientist takes a second step, an association by similarity. If, however, the mariner in turn through a new association by similarity applies the experience of the scientist to the directing of his boat, who will offer to decide as the relative intelligences involved?

What may from the standpoint of the materials used be spoken of as a highly developed mechanical intelligence may, from the standpoint of that which is done to and with the materials, be considered a highly developed abstract intelligence. It has been easier to develop tests of the use of language and of numbers and other symbols because they enter more commonly into the experiences of individuals, but the "thing-thinkers" also make their contribution to the sum total of human attainment. The simple relations of form, place and movement in the herewith described tests sample only some of the beginnings of the reactions of an individual to things about him. Our tests have still to show how far removed are the reactions of the individual to wind and tide and stars.

The points at issue may be illustrated by a consideration of the following statement of Terman:

"Many criticisms of the current methods of testing intelligence rest plainly on a psychology which fails to distinguish the levels of intellectual functioning or to assign to conceptual thinking the place that belongs to it in the hierarchy of intelligences. If an intelligence test can be shown to depend upon the language factor (i.e., upon the ability

---

usually far away from us. But the full consequences of the discovery were not realized at first. One such consequence is that the stars are not seen in their true places, that is, in the places which they truly held when the light left them (for what may have happened to them since, we do not know at all—they may have gone out or exploded). Our earth is only moving slowly compared with the great haste of light; but still she is moving, and consequently there is 'aberration'—a displacement due to the ratio of the two velocities, easy enough to recognize now, but so difficult to apprehend for the first time that Bradley spent two years in worrying over the conundrum presented by his observations before he thought of the solution. It came to him unexpectedly, as often happens in such cases. In his own words—'at last when he despaired of being able to account for the phenomena which he had observed, a satisfactory explanation of

to think in terms of symbols), this is sufficient in the eyes of some psychologists to condemn it as non-valid. The subject who cannot acquire a normal vocabulary, see the point of a fable, or be taught to read a paragraph with understanding is considered to have demonstrated his intelligence if he can trace a simple maze or assemble the fragments of a form board. In fact, the "idea thinker" is sometimes spoken of disparagingly or even a little contemptuously, particularly in the case of a child whose superior ability in this respect places him in a conspicuous contrast with other children of the same age.

But if intelligence is the ability to think in terms of abstract ideas, we should expect the most successful intelligence tests to be just those which involve the use of language or other symbols. We should also be justified in demanding that an intelligence test should correlate well with what we may call "school educability." As a matter of fact, it is precisely tests of this type which are surviving in the struggle for existence: tests involving arithmetical reasoning, language completion, naming opposites, matching proverbs, completing analogies, understanding difficult passages, etc. The list could of course be greatly extended, but all tests which are notably successful in measuring intellectual differences among adults have something in common with those named, even when they work with concrete materials instead of with words.

With this conception of intelligence we can understand why it is so difficult to devise tests of the non-verbal or 'performance' type which will bring out intellectual differences much above the level of the average child of ten or a dozen years// That it is difficult must be admitted by anyone who has attempted it and checked up his results. Non-verbal tests of the type used in the army Beta and performance examinations have of course proved useful on the intermediate levels of mental ability. However, even for children of eight or ten years, no purely non-verbal battery of tests yet devised is as successful as any one of several existing batteries which make some demands upon the language factor. Tests for very young children must of course be still simpler and work largely on the perceptual or at best on the lower representative and reproductive level. With very young infants, intelligence tests, as such, are of course not applicable. The best we can do here is to determine the presence of a nervous organization which will later make intelligence

---

them occurred to him all at once when he was not in search of it.' He accompanied a pleasure party in a sail upon the river Thames. The boat in which they were was provided with a mast which had a vane at the top of it. It blew a moderate wind, and the party sailed up and down the river for a considerable time. Dr. Bradley remarked that every time the boat put about, the vane at the top of the boat's mast shifted a little, as if there had been a slight change in the direction of the wind. The sailors told him that this was due to the change in the boat. not the wind; and at once the solution of his problem was suggested. The earth, running hither and thither round the sun, resembles the boat sailing up and down the river; and the apparent changes of wind correspond to the apparent changes in direction of the light of a star." From H. H. Turner's Introduction to "The Foundations of Einstein's Theory of Gravitation," by Erwin Freundlich. pp. 11, 12. Cambridge University Press 1920.

possible (e.g., tests of ability to acquire conditioned reflexes.)" (Intelligence and Its Measurement: A Symposium: § II by L. M. Terman. p. 129, The Journal of Educational Psychology, Vol. XII, No. 3, March 1921.)

It is as yet too early in the history of the development of intelligence tests to rest the argument on the "survival of the fittest," and the question is not, e.g., whether "a child cannot acquire a normal vocabulary" or is not possessed of "school educability," but whether he has had a fair chance and opportunity for such acquisitions so that his intelligence can be properly gauged by them, or whether his intellectual development has taken some other course and must therefore be tested by other means. Further, the above-quoted writer's notion of "conceptual thinking" must be regarded as faulty if it leads to an overemphasis of the seeming discreteness, to wit, "levels" and "hierarchies," in the development of intelligence rather than to a clear recognition of the essential sameness and continuity in the psychological character of the integration which we call intelligence. The concluding sentence of the above passage is especially open to criticism.

The *ability* to acquire conditioned reflexes can be evidenced only by their acquirement, and such acquirement may well be taken as marking the beginning of intelligence (and not simply of the presence of a nervous organization). There is good reason to believe that between this beginning and the most abstract ideas no qualitatively new processes intervene; the difference may be said to be quantitative rather than qualitative, i.e., more conditioned reflexes.* The *meaning* of the second or conditioning stimulus is found in the response which was made to the first, or putting it the other way around, the fact that the same response as was given to the first stimulus (or sensory pattern) is now given to the second contributes to the meaning or concept of the second situation. "In this sense," as James has said in a familiar passage, "creatures extremely low in the intellectual scale may have conception. All that is required is that they should recognize the same experience again. A polyp would be a conceptual thinker if a feeling of 'Hullo! thingumbob again!' ever flitted through its mind."† A succeeding passage of James in regard to the possibility of abstract ideas but apropos to the present discussion—is also worth recalling.

"But truly in comparison with the wonderful fact that our thoughts, however different otherwise, can still be of the same, the question whether that same be a single thing, a whole class of things, an

*See statement of writer, Journal of Educational Psychology, Vol. XII, No. 4; April 1921, pp. 210, 211.

†James: Psychology Briefer Course, p. 240.

abstract quality or something unimaginable, is an insignificant matter of detail. Our meanings are of singulars, particulars, indefinites, problematics, and universals, mixed together in every way. A singular individual is as much conceived when he is isolated and identified away from the rest of the world in my mind, as is the most rarefied and universally applicable quality he may possess—being, for example, when treated in the same way. From every point of view, the overwhelming and portentous character ascribed to universal conceptions is surprising. Why, from Socrates downwards, philosophers should have vied with each other in scorn of the knowledge of the particular, and in adoration of that of the general, is hard to understand, seeing that the more adorable knowledge ought to be that of the more adorable things, and that the things of worth are all concretes and singulars."*

The reason why it has been so difficult "to devise tests of the non-verbal or 'performance' type which will bring out intellectual differences much above the level of the average child of ten or a dozen years."† may be due to the fact that the verbalist and the scholastic have hitherto been the ones chiefly interested in the development of intelligence tests and they have naturally chosen tests in the use of which their own intellectual powers will not suffer by comparison. That is to say, the intellectual differences at issue and the methods of testing them may be traced a step further back to individual preferences (possibly in part innate, but more probably in large part the result of environment and training) for the various forms of sense material or knowledge, especially as between that of the, as often described, "higher" senses of sight and hearing and the "lower" muscular or kinaesthetic sense. It may well be questioned whether the kinaesthetic should be classed with the senses of smell and taste or whether it is a peer of sight and hearing in the intellectual development of some individuals. The education of the school would seem to depend chiefly on the last named, and when these possibilities are well considered, it may not appear even to the visualist and the audile absurd to argue that the above noted lack of differentiation among school children in the herewith described and similar tests after the ages of nine or ten is *in part* due to a retardation in the development of kinaesthetic knowledge.‡

*James: op. cit. p. 242.

†Before dismissing the possibilities of performance tests above the intellectual level of 10 or 12 year old children, the recent proposals of Stenquist (Teachers College, Col. Univ., Contributions to Education No. 130, 1923) and Kohs (Intelligence Measurement—A Psychological and Statistical Study based upon the Block Design Tests, MacMillan Company, 1923) especially merit consideration.

‡Most of us have sufficient kinaesthetic knowledge to devise tests up to this age. Series I of the writer's "Group Tests of Intelligence" and part of Series II are of this nature. Johnson and Schriefer (Jour. of Ed. Psych., Oct. 1922, p. 417) find a correlation of + .83 between Series I of these tests and the Pintner-Patter-

The writer will not pursue the argument further at this time, but commends to the reader's attention the excellent chapter of Pear on "The Intellectual Respectability of Muscular Skill"* from which the following suggestive passages are quoted.

"The word intellect has been used in many senses, of which at least two are important in this connection. It may cover all forms of cognitive events, in which case intellection becomes merely the equivalent of cognition. On the other hand, it may designate the processes of elaboration which the mind applies to its material, and in particular that kind of elaboration which leads to conceptual thinking. The point of view which this chapter seeks to justify is that, in both these senses of the word 'intellect,' muscular skill has a higher intellectual value than is usually assigned to it, and that this value is susceptible of being increased considerably.

We may first inquire whether such skill can justifiably be described as intellectual.

Bodily skill, or the ability to deal with the world by means of one's muscles, joints and tendons, carries with it a specific and unique kinaesthetic knowledge. This knowledge can obviously be called intellectual in the first sense mentioned above. Yet it is as incapable of perfect translation into the terms of another sense as is music into colors or words, for as we shall see, it possesses scarcely any words in its own right. Many persons therefore require to make a mental effort in order to realize that, in spite of its comparative wordlessness, kinaesthetic knowledge may form a basis for conceptual thinking, thereby establishing a primary claim to be called intellectual in the second sense of the term." (Op. cit. p. 208.)

"In a psychological discussion of this subject it is obviously relevant to comment upon the present disparagement, by an influential section of the community, of kinaesthetic knowledge. The tradition that knowledge worth having is almost exclusively confined to that which has reached us through our eyes and ears has been confirmed and hardened by the powerful mechanisms of class distinction and class tradition, and by generations of a certain type of teacher in school and university. To such influence the ordinary man owes his view of culture. From most people, therefore, the motiles, as compared with their socially-established brothers, the visiles and audiles, seldom get fair play. The two latter groups, themselves recognized as belonging to a kind of *intelligentsia*, not infrequently fall into the bad habit of regarding with

---

son Scale of Performance tests in the case of 88 children ranging in age from 3 years to 9.8 years, whereas the correlation between Series I and the Stanford Binet was + .55. Although the number of cases is very small, it is interesting to note that there is little or no improvement in the Pintner-Patterson score after the age of 7, although there is improvement in the Binet score. Is this the effect of schooling?

*Chapter XII of "Remembering and Forgetting," by T. H. Pear. E. P. Dutton & Co., New York; also published in substance in the British Journal of Psychology, 1921, XII, pp. 163-180.

some contempt all persons whose motor activity is expressed through channels other than the socially-approved ones of speech and writing. (Note: Mr. Bertrand Russell has placed on record his suspicion that what people should mean by intellect is simply 'certain habits in the use of words,' and his lack of 'mystical reverence for these habits.' ["Mind," XXIX, N.S. No. 116.]) A vicious circle is thus made; many intelligent persons are never encouraged to contemplate the study of such non-verbal occupations, and not a few are actively prevented from taking them up. This is one reason why the intelligent professional in sport, the intelligent handwork instructor, and the intelligent teacher of games is still so much of a rarity that when they do appear one never fails to remember them. If the knowledge obtainable through kinaesthesis were increased and transmitted to others in the best and quickest ways by teachers who were themselves good performers, their social status would rise. But this is the less important consideration; their intellectual status would be improved too. By increasing their kinaesthetic experience, they would not only have gained deeper knowledge of a comparatively new aspect of their world, but such knowledge might form basic material for the elaboration into concepts which characterizes intellect in the second sense of the term; the process by which we acquire the knowledge of truth as distinguished from the knowledge of facts." (Op. cit. pp. 229-230.)

"The relation of these considerations to the doctrine of *Bewusstseinslagen* or 'conscious attitudes' is obvious and interesting. (Note: Cf. E. B. Titchener, Experimental Psychology of the Thought Processes, pp. 98 f., 180 f.) It seems certain that the thinking which goes on in the possessors of a rich and pliant kinaesthetic memory tends to be strongly influenced by such conscious attitudes, perhaps to the discouragement of visual and auditory images. How far the left wing of "behaviorism' is constituted by psychologists with this type of mentality, whether, consciously or unconsciously, they contemn the visual-auditory intellect in much the same way that many intellectuals contemn the kinaesthetic variety; to what extent and in what ways these different kinds of mental apparatus may assist in the formation of those different attitudes towards life which are called introvert and extrovert respectively; (Note: Cf. C. G. Jung, Anayltical Psychology, pp. 287-98, and Psychologische Typen, Zurich, 1921) the rôles which different kinds of imagery play in aesthetic appreciation, in the causation and cure of mental disorders—all these problems are related to this subject, and study of them has scarcely yet begun." (Op. cit. pp. 230, 231.)

The above considerations in regard to the series of form board and performance tests described in the following sections, and the results of their use with school children as set forth in section 3, may be briefly summarized as follows:

It is found that the tests and methods of scoring presented in this monograph do not differentiate in the matters tested between children above the age of nine or ten years. The finding may depend on several factors:

(1) The methods of scoring may be in part at fault. Pertinent differences which the trained observer recognizes are not gauged merely by the time taken or the number of moves made by the subject.

(2) The tests may be too simple, and this is undoubtedly the chief reason for the finding. The fact that tests of this character which would be more discriminiative in the higher ages have either not been produced or have failed to secure general recognition may only argue that those whose experience or knowledge would most qualify them for the devising of such tests have not been as interested in the testing of intelligence as have those of a more verbal or scholastic turn of mind. The importance of kinaesthetic knowledge has been emphasized in the discussion, although the possible differences at issue are not solely kinaesthetic. If, in place of the teacher, the psychologist and the psychiatrist, who have been the creators of our present intelligence tests, a group of skilled artisans, engineers, higher mathematicians, physicists, astronomers, musicians, sculptors, surgeons, etc., should seriously undertake the devising of a series of intelligence tests, the narrowness and specialization of current methods of estimating intelligence might well be demonstrated.

(3 The fact that the learning of school children consists in good part in the forming of certain habits in the use of words and of numbers and in general of the symbols for things may result in a retardation in the development of other sorts of knowledge, as of things and their relations, which in some amounts to arrest and in others is only made up for in the practical pursuits of after school life. The bizarre performances of some children and adults with high intelligence ages and quotients (as determined by verbal and scholastic tests) in the "simple" tests of this series may thus be accounted for.

SECTION II—DESCRIPTIONS OF THE TESTS

1. COLOR-FORM TEST*

This test is patterned after material which is used in the Seguin room for the training of mental defectives at the School for the Feeble-minded at Waverley, Massachusetts. It consists of sixteen blocks, of which four are square, four triangular, four diamond shaped, and four circular. In addition, one block of each shape is painted a bright red, blue, green and yellow.

In giving the test the blocks are laid on a table so that no two blocks of the same shape or color are contiguous. The examiner then picks up a square, and, tracing the outline with his finger says, "FIND ALL THE BLOCKS WHICH ARE JUST THE SAME *SHAPE* AS THIS ONE." The block is held before the child an instant, then put back in place, and the time taken by the child to pick out the four squares is recorded by means of a stop-watch. Then the squares are replaced, and the examiner picks up a red block saying, "FIND ALL THE BLOCKS WHICH ARE THE SAME *COLOR* AS THIS ONE." Then the block is replaced, and the time is taken as before. The other tests are given in the same way. The same order is always followed, namely, square, red, triangle, blue, diamond, green, circle, yellow. It will be noticed that when the child is picking out forms he must disregard color and attend only to form, and vice-versa.

2. FORM BOARD 1A

This board was evolved for the purpose of providing a performance test which would serve to differentiate in the higher grades of defect as the Seguin board does in the lower. It consists of a board 17.5 by 11 inches in which four pairs of depressions, eight in all, are cut. These

---

*The Color Form Test, Form Boards, Nos. 1A and 1C, 2, 3, and 4, and the Furniture Test (test number 9 of this series) were first described in an article on "Form Board and Construction Tests of Mental Ability," by W. F. Dearborn, J. E. Anderson, and A. O. Christiansen, in the Journal of Educational Psychology, Vol. VII, Oct. 1916. The Color Form Test and Form Boards 3 and 4 were devised by the first named of these writers; tests 1A and 1C and 2 were devised by him in collaboration with Dr. J. E. Anderson; and test 5 (test 9 of the present series) in collaboration with Mr. A. O. Christiansen. For permission to reproduce the original charts of these tests (Figures 1, 2, 3, 4, and 5) the writers are indebted to Warwick and York, Inc.

Information in regard to the cost of the testing material described in this section may be secured by addressing the Secretary of the Psycho-Educational Clinic, Palfrey House, Oxford St., Cambridge, Mass.

depressions are the same shape as four of the blocks in the Seguin board, though they differ somewhat in size. (See Plate 1) One depression of each form and size may be filled by two blocks which are formed by cutting lengthwise a block which just fits into it; the other depression of the same shape may be filled by two blocks which are produced by cutting the fitted block crosswise.

In placing the form board before the subject, note that the long edge of the board and the rectangular depression are put nearest to him. The blocks should be arranged in four piles and placed within easy reach. Care should be taken that two blocks of the same shape do not come together in the same pile, and that the piles should be topped by blocks 4, 1, 8, and 5. An arrangement of the blocks in such an order is shown in the diagram of Plate 1, in which the bottom row (Nos. 3, 5, 7, 4), or the one nearest the board, represents the bottom of the piles, and the upper row (4, 1, 8, 5) the top of the piles.

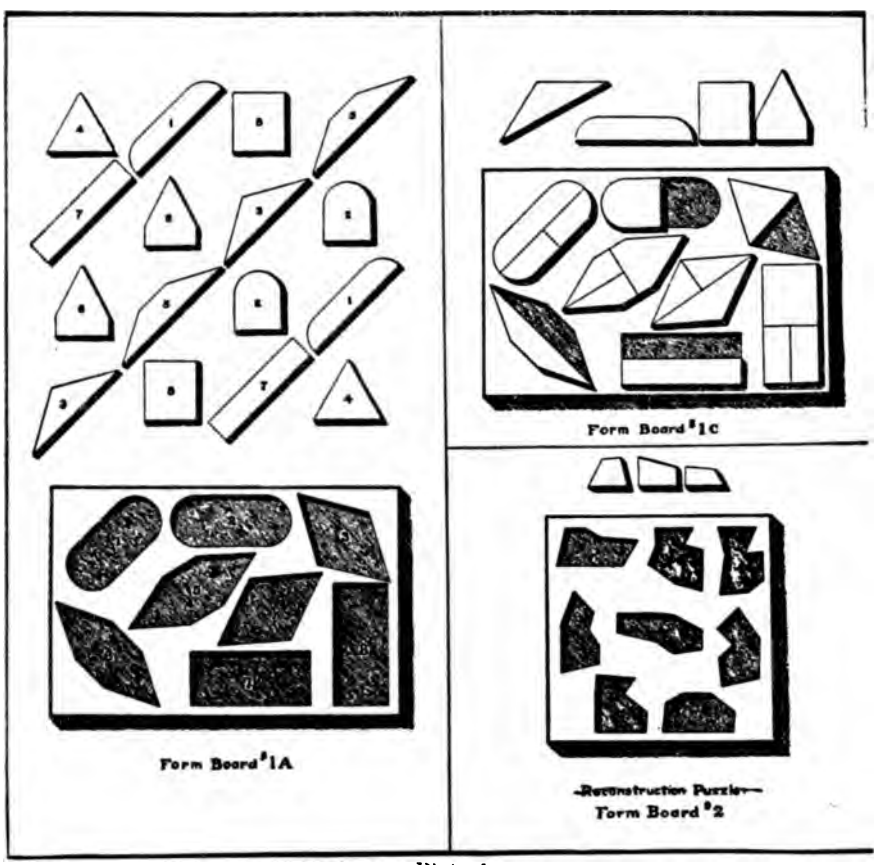


Plate 1

The examiner says to the subject, "HERE IS A BOARD WITH A LOT OF HOLES IN IT. DO YOU SEE THEM? (pointing to the depressions) AND HERE (pointing) ARE A LOT OF BLOCKS WHICH WILL FIT INTO THESE HOLES. THERE ARE TWO BLOCKS FOR EACH HOLE. I WANT TO SEE HOW FAST YOU CAN PUT THE RIGHT BLOCKS INTO THE RIGHT HOLES UNTIL ALL THE HOLES ARE FILLED. YOU MAY USE ONE OR BOTH HANDS. READY, GO!"

The test is scored (1) by the time and (2) by the number of moves made.

The examiner should take the time from the word "Go" until the last block is correctly placed. If the subject has not finished in five minutes, record the time as five minutes plus (5+) and mark "incomplete." Then show the subject how to complete the test.

Count the number and kind of moves according to the following directions:

A move is counted whenever the subject places or attempts to place a block in a depression. A move is counted as right when the block is fitted or tried in a depression in which it will fit, even although it is not a proper move for the completion of the test. It is counted as wrong when tried or placed in a depression in which it will not fit.

If the block tried will fit, it is counted as a right move even if the subject does not succeed in getting the block into the depression.

The same rules apply when a depression has been partially filled.

Notice that the removal of a block from a depression is not counted as a move. The move is counted only when the block is replaced on the board, or when an attempt is made to replace it.

If the subject fits two blocks together in his hand or on the table *two* moves must be counted when he places the blocks on the board.

The score is obtained by adding the right and wrong moves to get the total number.

3. FORM BOARD 1C

The board is the same as that used for 1A, but four of the half blocks are replaced by eight quarter blocks. The blocks which are not used should be put aside where the subject cannot see them. (See Plate 1)

The blocks are *first* arranged as shown in Figure 1 of Plate 2 with the four half blocks as indicated at the edge of the board farthest from the subject. A cardboard shield is then placed over the board so that only the two lower right hand depressions are visible, as shown in Figure 2. The long rectangular block is placed beside the board.

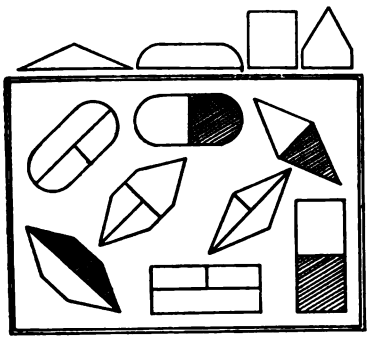


FIGURE 1

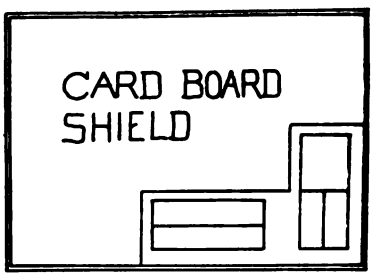


FIGURE 3

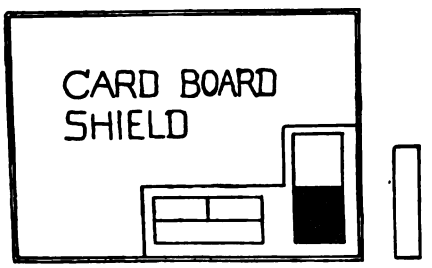


FIGURE 2

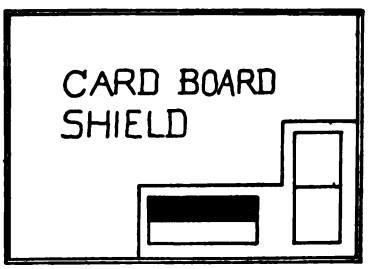


FIGURE 4

Plate 2

The examiner then says, "HERE IS A LONG BLOCK. I WANT YOU TO CHANGE THE OTHER BLOCKS AROUND IN THESE TWO HOLES SO THAT THE LONG BLOCK CAN BE FITTED IN AND BOTH THE HOLES WILL BE FILLED."

If the subject fails to solve the problem correctly in half a minute (30 seconds) he is shown how to do it. (See Figure 3.)

The board is now set up for the test by taking out the upper of the long rectangular blocks (see Figure 4 of Plate 2) and removing the cardboard shield. This leaves the board as shown in the diagram on Plate 1.

The instructions to the child are as follows: "NOW WHEN I SAY *GO* I WANT YOU TO FILL ALL THE HOLES AS QUICKLY AS YOU CAN WITH THESE FOUR BLOCKS (pointing to the four blocks at the upper edge of the board). YOU MAY CHANGE THE BLOCKS AROUND AS YOU NEED TO BUT DON'T CHANGE THEM ANY MORE THAN YOU HAVE TO. READY? *GO!*"

The test is scored for time and moves according to the same rules that are given for the scoring of board 1A.

4. FORM BOARD 2. RECONSTRUCTION PUZZLE

The board for this puzzle contains eight irregular depressions as shown in Plate 1. Each of these depressions may be filled by a certain combination of three blocks, two of which are in the form of trapeziums, while the third is a trapezoid. The sides of the blocks are in the relation of 2:3:4 with one extra side of 2 on one block, 3 on the second, and 4 on the third.* The board is placed on the table with the long axis of the center depression toward the subject.

The examiner picks up the three blocks and hands them to the subject, saying, "HERE ARE THREE BLOCKS WHICH FIT NICELY INTO EACH ONE OF THESE HOLES. WHEN I SAY *GO* I WANT YOU TO SEE HOW QUICKLY YOU CAN FILL THESE HOLES, BEGINNING WITH THIS ONE (pointing to Number 1). When the first depression is filled, point to hole 2 and say, "NOW PUT THE BLOCKS IN THIS ONE."

If the subject is not successful after three minutes' work at the first depression, the examiner should solve the problem, remarking how well the blocks fit. In subsequent problems if no successful placement is made within two minutes, the attempt is counted a failure. If one or two blocks have been correctly placed, however, a three minute limit is allowed.

The scoring of this board is on the basis of time only. Time is taken for each depression, and the total for all depressions is found.

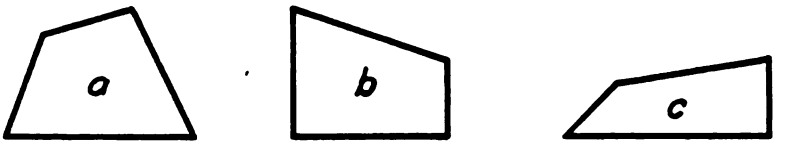
The board as described above is a modification of the first form in which this test was made.† The data obtained indicate that the test, even in the modified form, is much too difficult for the younger children. This is due, perhaps, to the fact that in order to fill some of the depressions the blocks have to be turned so that different faces are uppermost. To meet this difficulty a board was made in which there are but four depressions, each of which may be filled without turning any block. This board, however, has not yet received a thorough trial. (See Plate 3.)

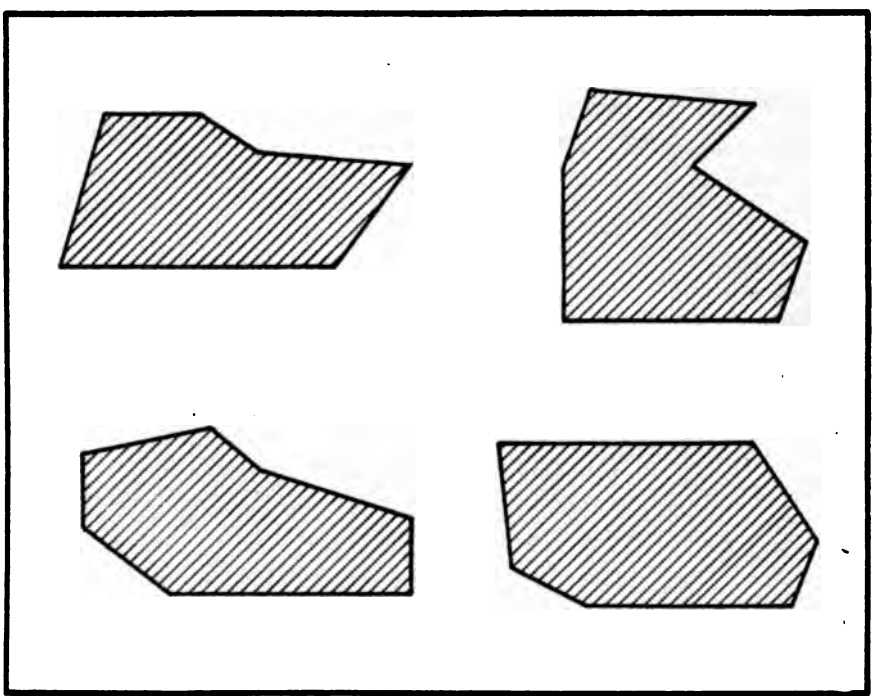
---

*These forms were adopted at the suggestion of Mr. Herbert Sturgis, then a graduate student in psychology at Harvard University.

†See the article of Dearborn, Anderson and Christiansen to which reference is made above.

Plate 3





Form Board No. 2.b

Shaded Areas = Depressions in Board.
"a" "b" "c" = blocks.

5. FORM BOARD 3

This board was first arranged by Dearborn in connection with a series of experiments on motor ability. The board as then used is shown in Plate 4.* This block test was later described to Drs. Healy and Fernald who adopted it, with modifications,† for use at the Juvenile Psychopathic Institute in Chicago.

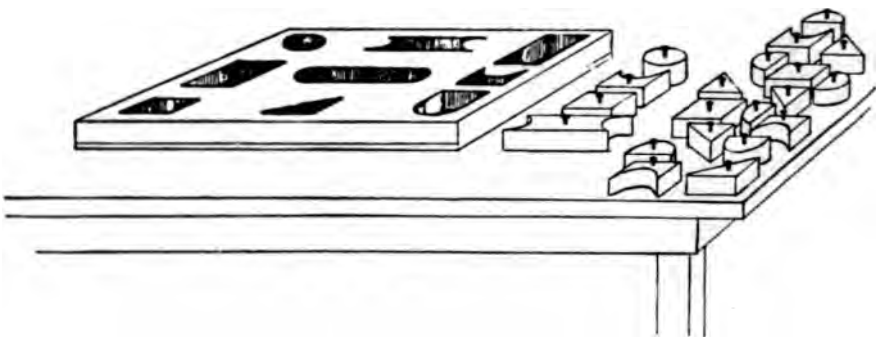


Plate 4

The Healy-Fernald method of using the board is the same as that described for 1A above. That is, all the blocks are taken out of the depressions, and the subject is asked to put them in as quickly as he can. This we believe to be undesirable, since several placements which are not easily recognizable as errors at the time they are made, nevertheless are errors, and must be corrected. Other placements, which to the superficial observer may appear no better, will lead directly to the solution. This introduces an element of chance into the problem, since by his first placements one subject may set up the board in such a fashion that he has a much harder problem than another individual of equal ability who starts in another way. This fact, it is believed, accounts in a great measure for the large variability which is found in the performance of the test, when given in this manner. The following method of using the board which was used in the original test, largely minimizes this difficulty, in that exactly the same problems are presented to each subject.

---

*This figure is reproduced from a theses by L. H. King on "The Validity of Motor Tests of Mental Efficiency," submitted for the degree of Master of Arts at the University of Wisconsin in 1906, and on file in the University Library.

†Psychological Review Monograph Series, Vol. XIII, No. 2, figure 8.

This form board was also employed as Test 5 of the "Performance Scale Examination" of the series of mental tests used in the U. S. Army during the war. The directions for giving and scoring the test in its use in the army were as follows, in which "E" stands for examiner and "S" for the subject being tested. (A problem similar to A was first given as a demonstration problem.)

(a) Examiner now presents the board arranged for problem A, saying, "WITHOUT MAKING ANY MORE MOVES THAN YOU HAVE TO, CHANGE THESE BLOCKS AROUND SO YOU CAN FIND A PLACE FOR THE EXTRA SQUARE (pointing to square). DON'T HAVE ANY BLOCKS LEFT OVER. READY—GO AHEAD."

(b) Examiner now presents the board arranged for problem B, saying, "I WANT YOU TO CHANGE THESE BLOCKS AROUND SO YOU CAN FIND PLACES FOR THESE TWO EXTRA SQUARES (pointing to them). READY—GO AHEAD."

(c) Examiner presents the board arranged for problem C, saying, "NOW I WANT YOU TO CHANGE THE BLOCKS AROUND SO YOU CAN FIND PLACES FOR THESE FOUR EXTRA BLOCKS. READY—GO AHEAD."

Examiner records the time in seconds from start to finish, and counts the number of moves. *A move is to be understood as placing or trying to place a block in some position on the board.* Taking a block out of position, and placing a block upon the table are not counted as moves.

Time for work on (a) and (b), *two minuutes each;* on (c), three *minutes.* If (a) is not solved in the time allowed, examiner demonstrates that correct solution before going on to (b).

Scoring.—If a problem is not solved within the time limit, score that part 0; but *if a correct solution has been accomplished,* give credit for time and for moves as follows:

MOVES				TIME		
(a)	(b)	(c)	Credit	(a) and (b)	(c)	Credit
..	....	8	5	0- 10	0- 20	5
..	....	9	4	11- 20	21- 40	4
3	5	10-11	3	21- 40	41- 70	3
4	6	12-14	2	41- 70	71-110	2
5-7	7-10	15-20	1	71-120	111-180	1

Note that the minimum number of moves for problems (a), (b), and (c) is 3, 5, and 8, respectively, and that the maximum raw scores are 8, 8, and 10, or a total of 20 points.*

In the present use of the test the procedure of giving the test is similar to the above, but the method of scoring is different.

The arrangement of the blocks and board for four problems are shown in Plate 5.

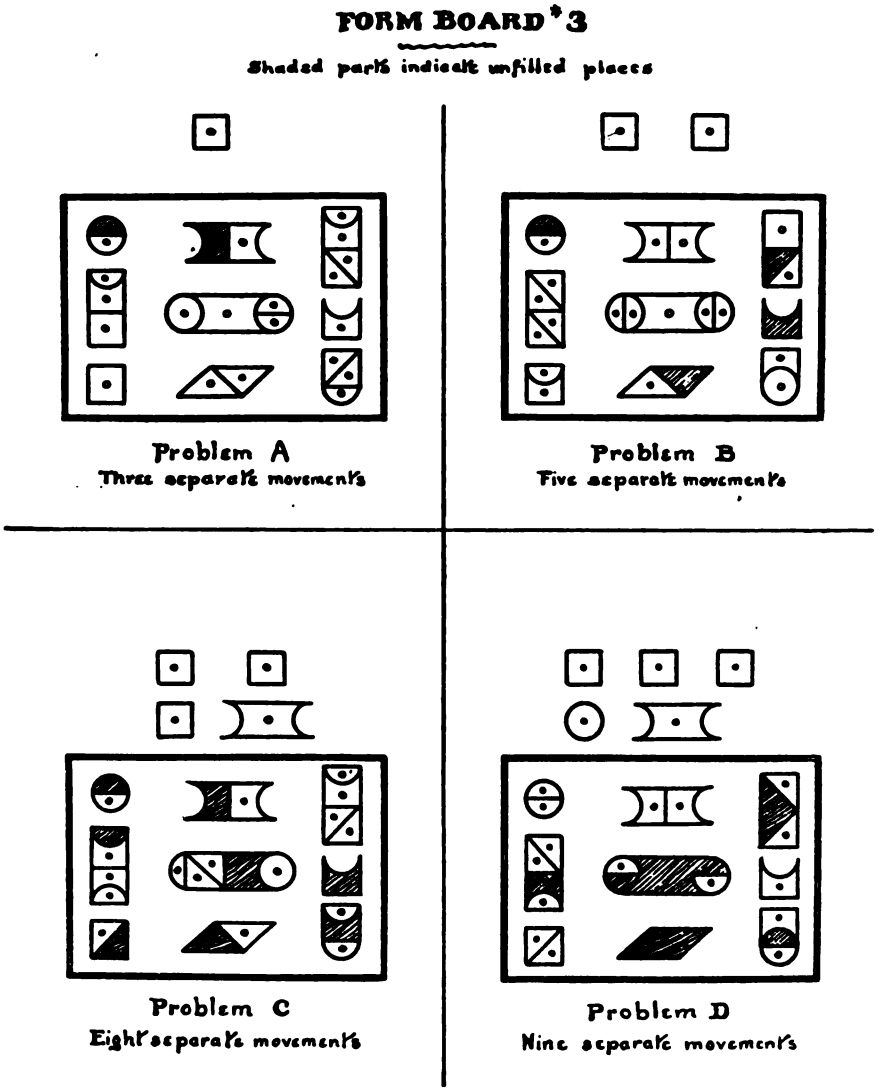


Plate 5

*The above statement is quoted from "Psychological Examinings in the United States Army," edited by Robert M. Yerkes, Memoirs' National Academy of Sciences, Vol. XV, p. 185. See also Army Mental Tests, by Yorkum and Yerkes, pp. 108, 109, Holt & Co., New York.

Problem A is used as an example. After the blocks are arranged, the examiner says to the subject, "I AM GOING TO SHOW YOU HOW TO DO THIS PUZZLE, AND THEN I WILL GIVE YOU ANOTHER LIKE IT TO DO YOURSELF. DO YOU SEE THESE HOLES OR SPACES IN THE BOARD (pointing to them) WHICH HAVE NO BLOCKS IN THEM, AND THIS SQUARE BLOCK WHICH HAS BEEN LEFT OUT? WHAT WE WANT TO DO IS TO MAKE A SPACE OR HOLE FOR THIS SQUARE BLOCK AND PUT THE BLOCK INTO IT IN SUCH A WAY THAT ALL THE HOLES IN THE BOARD WILL BE FILLED AND NO BLOCK WILL BE LEFT OUT. IN ORDER TO DO THIS WE MAY CHANGE THE OTHER BLOCKS ABOUT; ONLY WE MUST NOT MAKE ANY MORE CHANGES THAN ARE NECESSARY." The examiner then illustrates by filling the two empty spaces so as to clear a square space, in which he then places the square block.

The board is then concealed from the subject and set up for problem B as shown in the diagram. When it is ready the examiner presents it to the subject and says, "I WANT YOU TO PLACE THESE TWO SQUARES IN THE BOARD SO THAT ALL THE HOLES WILL BE FILLED AND NO BLOCKS BE LEFT OUT. MAKE JUST AS FEW MOVES OF THE OTHER BLOCKS AS YOU CAN, AND WORK QUICKLY."

Problem D is next presented in the same way, and problem C is given last, since it has proven more difficult than D.

Each problem is scored for both time and moves. If a problem is not solved in three minutes, it is recorded as "Incomplete," and the examiner passes to the next problem. As in the other tests, it is considered better to help the failing subject to find the solution quickly after the time limit is up rather than to break off suddenly, and thus emphasize his failure.

A move is counted whenever the subject places or attempts to place a block in an aperture of the board. If the block is one which will fit in the depression, that is, if there is no error of form, the move is counted as correct, no matter if it does not contribute directly to the solution of the problem. The removal of a block from the board does not count as a move, since the move is made when the block is replaced on the board. A move is recorded as incorrect or wrong when the subject attempts to put a block into a depression which it will not fit. Right and wrong moves are added to get the total.

6. FORM BOARD 4. TRIANGLE PERFORMANCE TEST

Formerly no board was used in giving this test. A chart containing 12 rectilinear figures was placed before the subject, and he was

given two blocks in the shape of right angled triangles. Then he was asked to make the figures on the chart with the two blocks. (See Plate 6.) The figures on the chart were smaller in proportion than the blocks, so that the problems could not be solved by direct super-imposi-

A "TRIANGLE" PERFORMANCE TEST

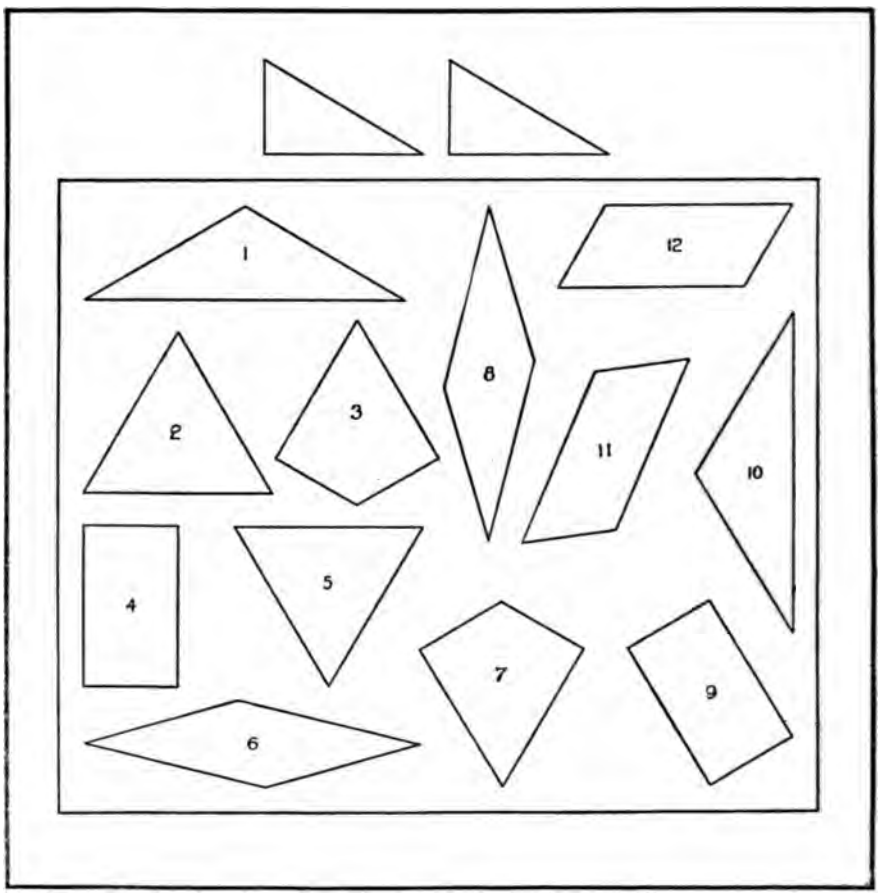
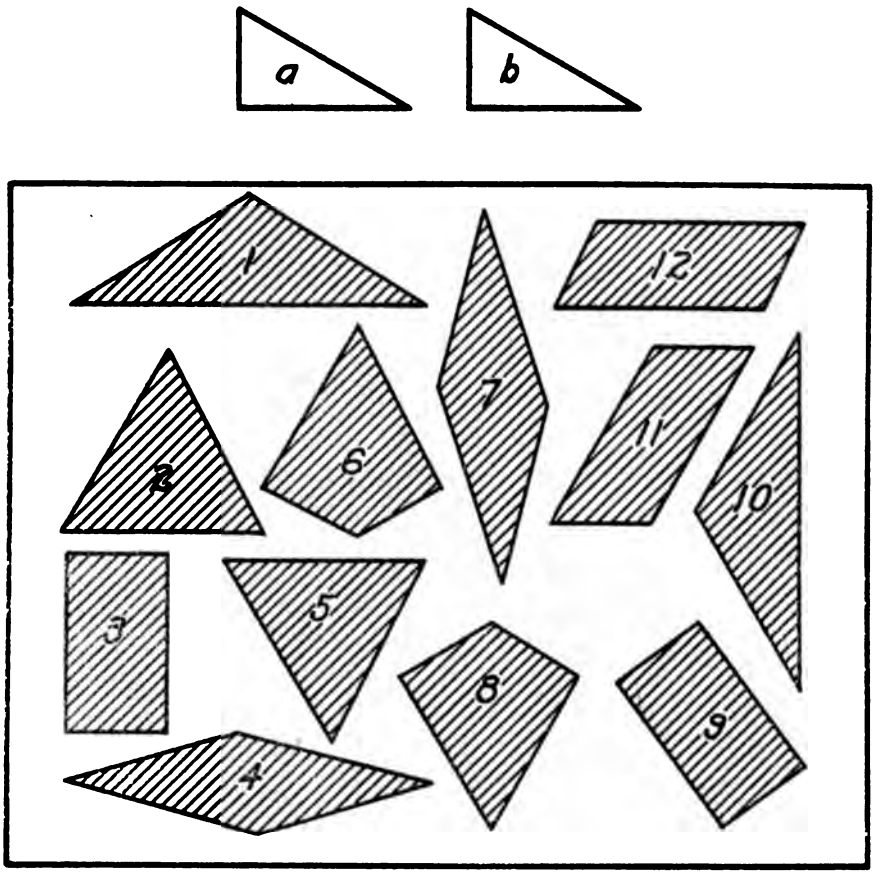


Plate 6

tion. This made the test a rather difficult one, and so boards were made in which depressions were cut, the same shape as in the original chart, but large enough so that the triangular blocks would fit. Plate 7 shows the board and blocks.

In giving the test the two triangular blocks are placed at the top of the board as shown in Plate 7, and the examiner says to the subject, "EACH OF THESE HOLES (pointing) CAN BE FILLED WITH THESE LITTLE BLOCKS. Illustrate by filling hole Number 1,

Plate 7.



Form Board No. 4.

Shaded Areas = Depressions in Board.
"a" and "b" = Triangular Blocks.

and replace the blocks in the same position at the top of the board.) NOW WHEN I SAY *GO* I WANT TO SEE HOW QUICKLY YOU CAN PUT THE BLOCKS INTO THE FIRST HOLE. READY, GO!"

If the subject is not successful within one minute, the examiner again shows how the problem is solved.

When the blocks are in place in the first depression the examiner says, "NOW, WHEN I SAY *GO* I WANT YOU TO TAKE THE

BLOCKS OUT OF HOLE NO. 1 AND PUT THEM IN HOLE NO. 2 (pointing); THEN TAKE THEM OUT OF NO. 2 AND PUT THEM INTO NO. 3, AND SO ON. READY, GO!"

As soon as each subsequent hole is completed the examiner says, "NOW TRY THE NEXT HOLE, NO. ..." This direction should be accompanied by pointing, and should be given quickly in order that the subject may lose no time.

If hole No. 2 is not filled in one minute the subject is quickly shown the solution, and is then asked to try No. 3. No help is given on any of the holes after No. 2. If the subject appears to give up because of his failure, he may be encouraged by the general direction, "TRY THEM SOME OTHER WAY. TURN THEM AROUND AND TURN THEM OVER."

If the subject tries to match the block in the air before putting them into the depressions the examiner says to him, "PUT ONE BLOCK IN FIRST AND THEN THE OTHER BLOCK."

The score is the total time taken from the word "Go" on depression No. 2 to the completion of No. 12, including the time of demonstration, if made, of the method of filling No. 2. If all the problems are not completed in ten minutes, record the time as ten minutes plus (10+) and mark incomplete. The subject may then be helped to complete the test in order to avoid discouragement.

7. LINCOLN HOLLOW SQUARE

This board was recently devised and added to the series primarily for the purpose of testing very young children. A second consideration was the desirability of a test which should be lighter and less bulky than many of the other form boards.

In the general principle underlying it, this board resembles a number of others in which the task is to fill various depressions with two or more blocks. It is, however, better fitted for use with young children than the Healy-Fernald Construction Puzzle A,*. and somewhat more convenient than Miss Kent's series.† In common with the other puzzle tests of the same nature, an element of chance enters into the solution of this board. How much this element affects the reliability of the test it is impossible to say. An attempt is made to obviate this difficulty, in that the child is given eight problems to solve instead of only one.

*Healy and Fernald, Tests for Practical Mental Classification. Psychological Monographs, Vol. xiii, No. 2. Also Pintner and Patterson, A Scale of Performance Tests, p. 44.

†Kent, G. H., A Graded Series of Geometrical Puzzles, Journal of Experimental Psychology, Vol. i, 1916, p. 40ff.

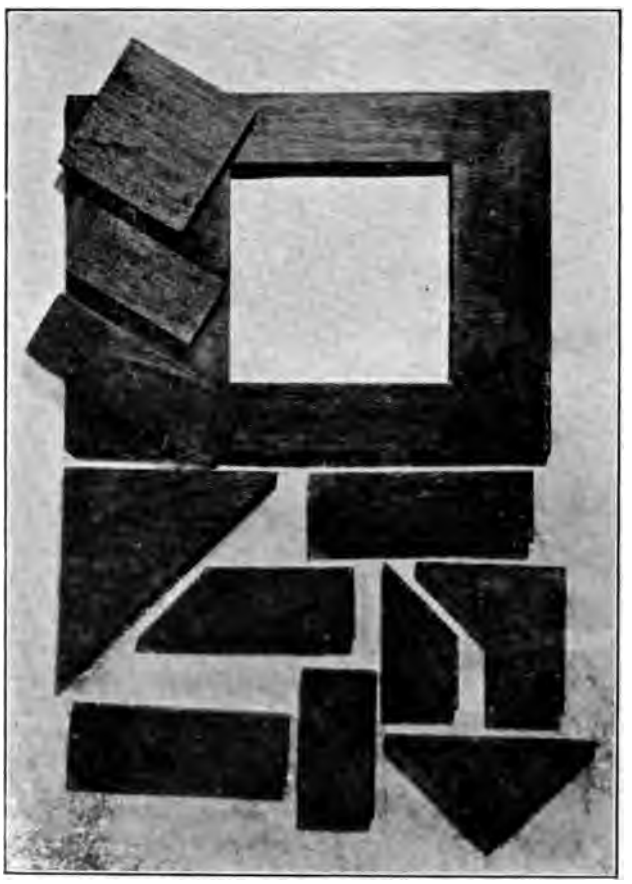


Plate 8

The board and eleven pieces, combinations of which will fill the square, are shown in Plate 8. The hole cut in the board is four inches square. Each piece is numbered inconspicuously on each of its larger surfaces, duplicate pieces having the same number. On the board are stamped, also inconspicuously, the numbers of the pieces which make up the eight combinations for the different problems.

The board is placed on a table or other flat surface, in such a position that the number combinations can be read by the examiner. The pieces are laid out at one side within easy reach, and arranged in numerical order so that any desired piece may be quickly found.

The examiner picks up the three No. 1 pieces and says to the subject, "DO YOU SEE THE HOLE IN THIS BOARD? (pointing) WELL, I WANT YOU TO FILL IT UP WITH THESE BLOCKS (handing them to him). DO IT AS QUICKLY AS YOU CAN. READY, GO!"

This problem is a very easy one, and was so devised that only a very low grade subject can fail in the solution. However, in case the subject is not successful within a minute, he is shown the solution, and the next problem is presented to him. This plan is followed with all the problems. When a subject wishes to give up before the expiration of the time limit, he should be encouraged by the direction "TRY THEM SOME OTHER WAY."

Time may be saved if the examiner will pick out the pieces for the second problem as soon as the subject starts working on the first, and so on. At the completion of each problem the board is lifted and the blocks are swept aside, thus quickly preparing for the succeeding problem.

The time is taken from the word *GO* until the problem is completed or until one minute has elapsed.

The score is the number of problems which are correctly solved within the time limit of one minute each.

8. SHAW PICTURE PUZZLES*

In undertaking the development of these puzzles the object was to produce a picture completion test in which the difficulties in the previous tests of the same nature might be avoided. It was in the first place, desirable to find a picture characterized by a unity not found in

*This test was first described by E. A. Shaw in the Journal of Applied Psychology, Vol. 2, No. 4, pp. 355-65.

tests of a like nature. Secondly, the effort was made to produce a test which would be superior in its mechanical aspects to any which had at that time appeared.

The Drug Store puzzle, which was the first of the series to be developed, is shown in Plate 9. The picture is mounted on a 16x12 board in which there is a drawer for keeping the insets.

There are ten situations depicted, each of which involves an action which is peculiar to a drug store. An essential part of each situation has been cut out and placed upon an inset. The depressions and insets are circular, and all are of the same diameter, thus avoiding the misconception involved when children simply try to fill the depressions rather than to complete the situations. The inset blocks are slightly thicker than the depth of the depressions so that the removal of the inset is facilitated.

In addition to the ten insets which complete the picture correctly, there are ten others which have some of the elements of the correct situations, but which do not offer such satisfactory completions as the first set. There are also ten insets which show articles that have little or no connection with a drug store, and ten blanks. (See Plate 10.)

In giving the test the board is placed on a table before the subject, and the drawer containing the insets is put at his right. The examiner says, "THIS PICTURE IS NOT FINISHED, SOME PARTS HAVE BEEN LEFT OUT, AS YOU SEE (pointing). YOU WILL FIND THE MISSING PARTS THERE (pointing to open box). I WANT YOU TO FIND THE RIGHT BLOCK FOR EACH HOLE AND PUT IT IN. OF COURSE, YOU CAN'T USE ALL THE BLOCKS, BUT YOU MAY BE SURE THAT THE RIGHT BLOCK FOR EACH HOLE IS IN THE BOX. YOU MAY USE ONE OR BOTH HANDS, AND CHANGE THE BLOCKS AROUND AS MUCH AS YOU WISH. WORK AS QUICKLY AS YOU CAN, AND LET ME KNOW WHEN YOU HAVE FINISHED IT. DO YOU UNDERSTAND? ALL RIGHT, GO AHEAD."

Time is taken from the examiner's "Go ahead" to the time when the individual indicates that he has finished. Filling all the holes does not constitute the completion of the task, since the subject may wish to make changes after his first attempts.

It is well to record the subject's general plan of attack, his method of work, and so on, for although the test is not scored on the basis of these things, the facts are often very valuable in making a final judgment on the case. If the subject has not mentioned the name of the picture during his work, he is asked at the end. "WHAT IS THIS A

Plate 9

Plate 10

PICTURE OF? WHY DO YOU THINK THAT? WHAT SEASON OF THE YEAR IS IT? WHY DO YOU THINK SO?" The answers to these questions are recorded.

The scoring is done on the following basis:

10 for each hole correctly filled.

5 for each hole filled with a second choice block.

3 for a first or second choice in the wrong hole.

0 for a hole left empty.

Minus 2 for a blank block in any hole.

Minus 3 for any block which has no connection with a drug store.

Thus the score ranges from minus 30 to plus 100 for a perfect performance.

9. FURNITURE TEST

In the article already cited* Dearborn and his co-workers describe a chair construction test in which the parts of a small rocking chair are presented to the subject (see Plate 11) with the instructions, "SEE IF YOU CAN MAKE A PIECE OF FURNITURE OUT OF THESE

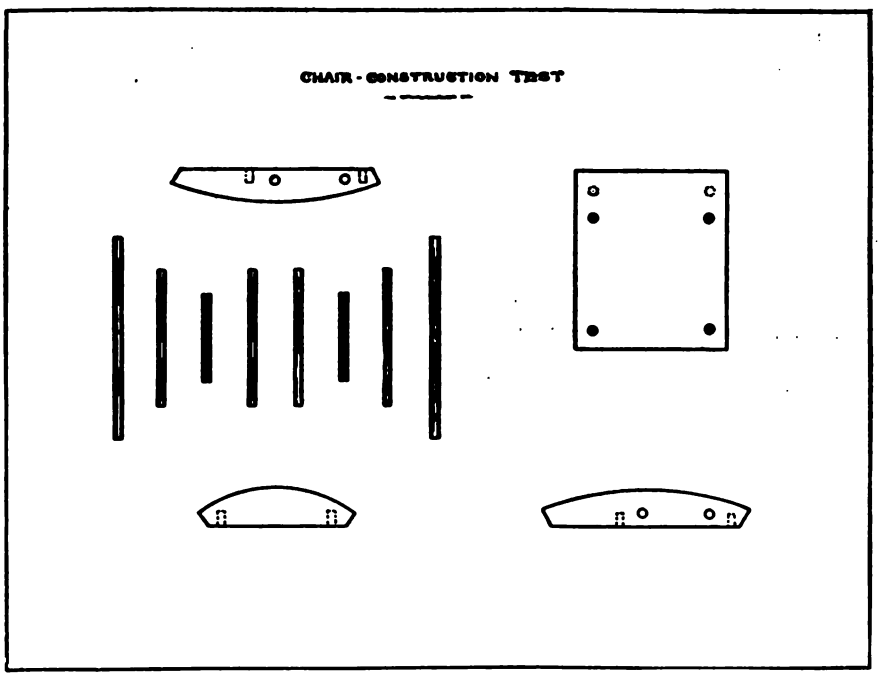


Plate 11

*Journal of Educational Psychology, October 1916.

BLOCKS. YOU MUST USE ALL THE PIECES." This test was elaborated later by Shaw* to make the test as described in the following paragraphs.

The apparatus for this test consists of 15 pieces of wood, as shown in Plate 12, out of which several pieces of furniture can be made. Two of these pieces of furniture, one a bed and the other a cradle, require all the blocks for their construction. One of the models is shown in Plate 13.

In giving the test the blocks are presented to subject as shown in Plate 12. Then the examiner says, "WITH THESE PIECES I SHOULD LIKE TO HAVE YOU MAKE A PIECE OF FURNITURE OR SOMETHING OF THAT KIND, USING ALL THE BLOCKS. YOU MAY MOVE THE PIECES AROUND IN ANY WAY YOU WISH. WHEN YOU HAVE MADE SOMETHING, TELL ME WHAT IT IS. WORK AS QUICKLY AS POSSIBLE."

If the subject makes something which does not require all the blocks, he is asked to try again, with a repetition of the injunction to use all the blocks. A third trial is allowed if the second is also unsuccessful.

The time for each trial is recorded, as well as the name given by the child to each object which he makes. A score for each part is obtained by allowing one point for every piece which is used in building. This gives a possible maximum score of 15 when all the blocks are used. One point is subtracted for each of the long or short rails which is not parallel with the large baseboard. The subject is graded on his best score in the three trials.

---

*The Constructive Type of Intelligence Test, E. A. Shaw. An unpublished thesis on file at the Graduate School of Education at Harvard University.

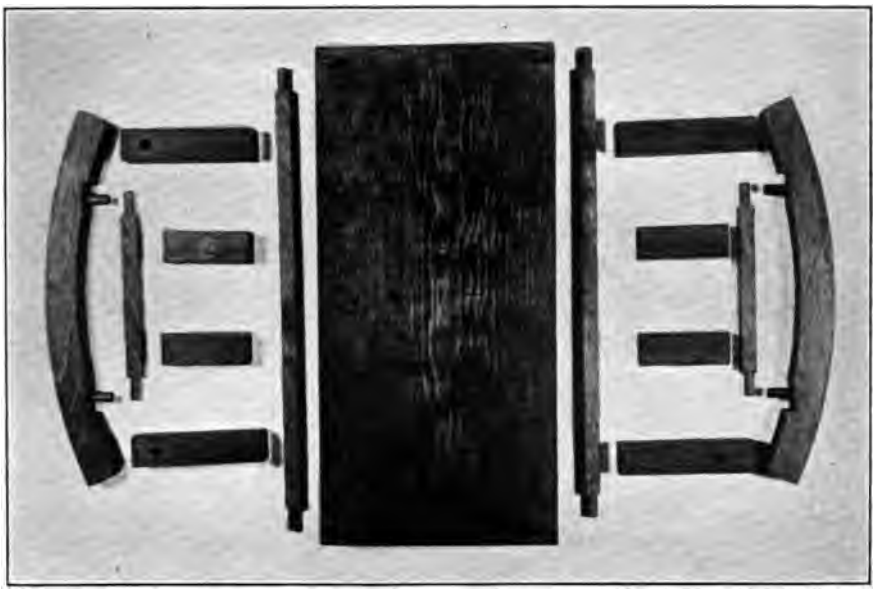


Plate 12

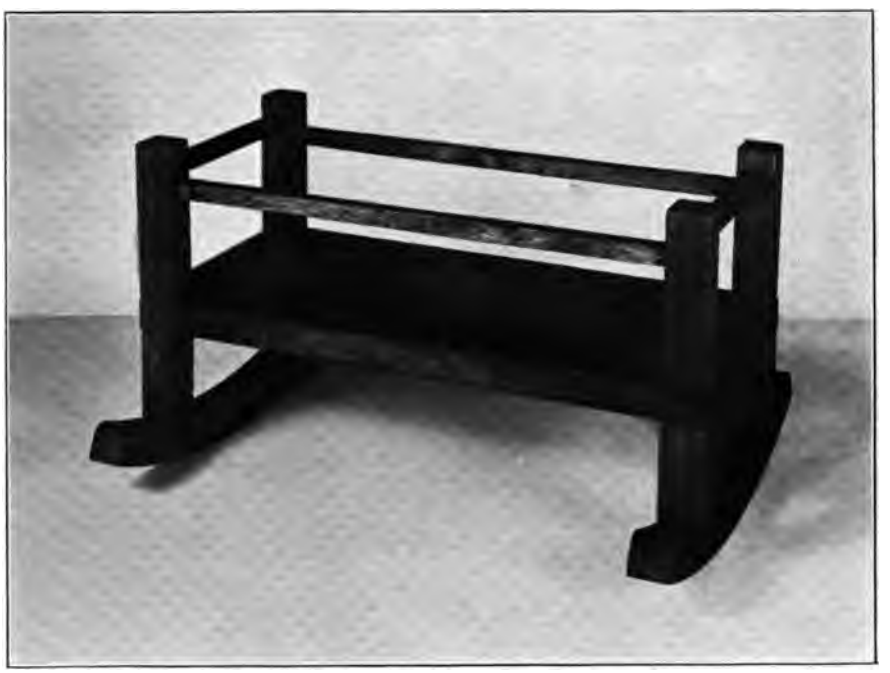


Plate 13

SECTION III—STANDARDIZATION OF THE TESTS

There are two processes necessary in the standardization of a test. In the first place, the method of giving the test must be worked out in some detail with due attention to clearness of directions, interest to the subjects, ease of manipulation, objectivity of scoring and similar factors. This phase of the standardization has been treated in the first section, where directions for giving and scoring the tests were presented. These methods of procedure should be followed in detail whenever the tests are given.

The second problem of standardization is that of determining norms or standards usually for various age groups, in order that the performances of an individual may be used as a means of classification. This involves giving the tests to relatively large numbers of individuals which are not selected in any way. It is a considerable task to do this with an individual test, first, because of the time which must be given to each subject who is examined, and also because of the difficulty of getting a group of subjects in which it is certain that there is no factor of selection operating. It should be pointed out in this connection that the mere piling up of large numbers of cases does not automatically insure correct or reliable standards. Many tests, both group and individual, lay their chief claim to attention in the fact that they have been given to many children. However, as Pintner and Patterson have pointed out, it is not necessarily true that a norm based on 200 or 2000 cases is more accurate than one based on 100. "As a matter of fact," they say, "the accumulation of an additional thousand cases to the first thousand or an additional hundred to the first hundred, may be simply a waste of time. The question (of accuracy of norms) resolves itself into a consideration of the group of individuals tested, the variation of the norm with the addition of each group of results, and the type of standardization required."*

Several methods of sampling have been tried for the purpose of getting groups of children in which no selective factors should enter. Terman, in getting data for the Stanford Revision of the Binet scale, selected "a school in a community of average social status, a school attended by all or practically all the children in the district where it was located. . . . To avoid accidental selection, *all* the children within two months of a birthday were tested, in whatever grade enrolled. Tests of foreign born children, however, were eliminated in the treatment of the

---

*Pintner and Patterson, A Scale of Performance Tests. Appleton, p. 72.

results."* In standardizing the Witmer form board, Young took two schools, one in a better and one in a poorer environment.† Simpson chose two groups of individuals, one known to be good, the other considered poor. He argued that the median performance of these groups would give a fair norm for the population at large, and as the individuals were rather carefully selected, there seems to be a justification for this claim.‡ Pintner and Patterson used two schools attended by children of the middle classes, and point out that this method probably gives reliable central tendencies, although there may be some inaccuracies in the upper and lower ends of their distributions.§ The Dearborn Intelligence Examinations were standardized by testing *all* the children in three "typical" communities.¶ This latter method is more desirable than any of the others, but it is practically impossible to standardize an individual test in this way because of the tremendous amount of time which would be required to obtain the data.

In gathering the data for the tests in this series it has not been possible up to the present time to carry out any one of the plans outlined in the foregoing paragraph. The results have come from many classes, schools and systems. Some of them were obtained in the preparation of doctorate theses, others were collected in the classes for training psychological examiners. The number of children given each test and the grades from which the cases were taken were both largely matters of chance. It is believed, however, that no very strong selective factor was at work in any school or grade where the tests were given. No results were used which were obtained from children tested because of scholarship or disciplinary difficulties.

COLOR FORM TEST

The results obtained with the Color Form test are set forth in tables 1 and 2.

Table 1. Summary of Color Form test results, first trial. Time in seconds.

---

*Terman, L. M., The Measurement of Intelligence. Houghton Mifflin, p. 52.

†Young, H. H., The Witmer Form Board. Psychological Clinic. Vol. X, 1916, No. 4, p. 93.

‡Simpson, B. R., Correlations of Mental Abilities. Teachers College, Columbia University, Contributions to Education, No. 3, 1912, p. 122.

§Pintner and Patterson, A Scale of Performance Tests. Appleton, p. 75.

¶Dearborn and Lincoln, How the Dearborn Intelligence Examination Standards Were Obtained. Journal of Educational Psychology, Vol. XIII, 1922, No. 7, p. 295.

TABLE 1

Age	Cases	Sq.	Red	Tri.	Blue	Diam.	Gr.	Cir.	Yel.	Total Time	Gen. Av.
4-5	10	11.4	6.9	11.4	6.0	12.5	6.6	7.3	4.6	75	8.2
6	7	14.7	6.8	7.8	4.5	9.0	4.1	7.0	3.5	54	6.1
7-9	3	7.0	7.0	5.6	3.3	4.6	4.0	3.6	3.3	24	5.2
All	20	11.9	6.9	9.3	5.1	10.4	5.3	6.5	4.0	153	7.4

Table 2. Summary of Color Form test results, later trials. Time in seconds.

TABLE 2

Age	Cases	Sq.	Red	Tri.	Blue	Diam.	Gr.	Cir.	Yel.	Total Time
				MEDIAN TIME						
3-11	1	x	10.4	x	7.0	x	8.0	x	11.0	
4-9	5	19.1	8.0	16.0	4.0	26.1	9.3	6.1	8.0	102.8
5-9	7	9.1	8.0	8.4	5.0	9.4	4.3	6.0	4.0	79.2
6-4	7	4.3	3.2	5.1	4.0	6.1	5.0	4.2	3.1	38.2
7-6	7	7.4	3.0	5.2	3.3	9.0	4.0	5.0	2.0	34.6
8-6	8	4.2	3.0	6.7	3.3	4.6	4.0	2.3	2.3	33.8
9-9	1	3.0	5.4	5.0	2.4	5.4	3.0	2.2	2.1	28.5
10-3	2	3.2	2.5	3.1	3.7	4.0	2.6	2.6	1.6	23.2
11-0	1	3.0	4.0	4.4	4.0	3.4	4.0	2.3	2.0	27.1

The figures in these tables show median times for the given number of cases except that the mean was taken for the two 10-year-old cases in Table 2. The 3-year-old was not able to pick out any of the forms.

Each set of data shows that the ability of children to pick up blocks increases with age. The variabilities were in some instances rather large, but their true significance cannot be gauged when the number of cases is so small. The indications are that the test will be useful for the third, fourth, fifth and sixth years. Possibly it will serve for differentiation a year or two further.

It will be noticed that the colors are quite uniformly picked up more rapidly than the forms. As the age increases the time for picking up forms decreases somewhat more rapidly than the time for picking up the colors, and in the upper years the differences are not so great. This seems to indicate that the sense of color is developed earlier than the sense of form, at least as far as the colors and forms in this particular test are concerned.

Although it is impossible to set up definite final standards on the basis of the data which are now at hand, it seems probable that the total time will be found the most convenient way of scoring this test. On this basis the following norms are suggested for tentative use.

hand, it is not possible to set up age standards on the basis of the number of moves. It seems worth while, however, to continue to record the number of moves and the number of errors, for even though these records do not contribute to the determination of a mental age they are certainly valuable in helping the examiner to form an estimate of the subject's general ability to work with his hands. This is extremely important in recommending proper treatment for the case.

There is considerable differentiation on the basis of time. It seems altogether probable that further results will show that the time for the 9 and 10 year groups should be decreased.

On the basis of the above results the following Tentative Standards for Form Board 1A norms are suggested.

Age	5	5½	6	6½	7	7½	8	8½	9
Seconds	210	200	195	190	165	135	125	120	115

To the performance age thus found on the basis of time taken, add one year; if the task is completed in from 16 to 18 moves, subtract one year if 24 or more moves are made.

FORM BOARD 1C

Tables 5 and 6 present the data gathered concerning Form Board 1C.

TABLE 4

Form Board 1A—Distributions of Time by Ages

Seconds	Age 5	6	7	8	9	10
0-9						
10-19						
20-29			1			
30-39			1			
40-49			1	1	1	
50-59		1	2	4	2	
60-69	1	4	12	11	1	
70-79		2	11	7	6	
80-89		0	12	11	4	4
90-99		1	17	4	3	2
100-109		2	14	5	3	2
110-119		1	5	10	2	0
120-129	1	2	11	4	4	1
130-139	1	3	8	3	3	1
140-149	3	6	8	9	2	0
150-159	2	3	4	3	1	1
160-169	1	1	11	2	1	2
170-179	0	2	3	2	1	
180-189	2	4	4	5	0	
190-199	2	2	10	7	1	
200-209		4	2	2	3	
210-219		3	3	0	0	
220-229		2	5	0	0	
230-239		2	1	1	1	
240-249	2	2	3	0	0	1
250-259	1	2	4	2	0	
260-269	2	2	5	2	0	1
270-279	1	0	3	1	0	
280-289		3	2	1	1	
290-299		1	1		1	
300 plus			2		0	
Incomplete	6	8	13	8	4	1
No. Cases	25	63	179	105	45	16
Median	197.5	189.0	133.1	119.5	121.2	115.0
Median age	5.71	6.51	7.57	8.43	9.51	10.42

It will be noticed that there is no significant difference between the median number of moves taken to solve the problem. This means that unless subsequent results are decidedly different from those now at

hand, it is not possible to set up age standards on the basis of the number of moves. It seems worth while, however, to continue to record the number of moves and the number of errors, for even though these records do not contribute to the determination of a mental age they are certainly valuable in helping the examiner to form an estimate of the subject's general ability to work with his hands. This is extremely important in recommending proper treatment for the case.

There is considerable differentiation on the basis of time. It seems altogether probable that further results will show that the time for the 9 and 10 year groups should be decreased.

On the basis of the above results the following Tentative Standards for Form Board 1A norms are suggested.

Age	5	5½	6	6½	7	7½	8	8½	9
Seconds	210	200	195	190	165	135	125	120	115

To the performance age thus found on the basis of time taken, add one year; if the task is completed in from 16 to 18 moves, subtract one year if 24 or more moves are made.

FORM BOARD 1C

Tables 5 and 6 present the data gathered concerning Form Board 1C.

TABLE 5

Form Board 1C—Distributions of Moves by Ages

Moves	Age 5	6	7	8	9	10
12	1	5	20	6	7	1
13		5	13	9	0	0
14		4	15	10	4	0
15	1	4	14	6	2	3
16	1	1	7	8	1	0
17		1	10	3	2	2
18		2	6	7	6	0
19		2	5	3	1	1
20	1	2	6	1	2	1
21		1	7	3	1	0
22		1	6	0	3	0
23	1	0	5	2	1	0
24	1	2	1	0	1	0
25	0	1	2	2	2	0
26	1	1	1	1	0	0
27	1		2	0	1	1
28			1	0	0	0
29			4	0	1	0
30			0	0	0	0
30 plus			3	1	2	1
Incomplete	17	31	51	43	8	6
Total	25	63	179	105	45	16
Median	Inc.	26.5	19.9	20.5	19.5	
Median age	5.71	6.51	7.57	8.43	9.51	

TABLE 6

Form Board 1C—Distributions of Time by Ages

Seconds	Age 5	6	7	8	9	10
0-9						
10-19						
20-29						
30-39						
40-49			1			
50-59			1			
60-69		1	0	2	1	1
70-79		1	2	2	0	
80-89			1	4	1	
90-99			5	2	1	
100-109		1	4	3	0	1
110-119			1	2	1	
120-129		1	7	2	3	
130-139		1	9	4	2	
140-149			1	3	2	
150-159		2	2	2	3	
160-169			4	3	2	
170-179		3	6	1	1	1
180-189	3	3	5	2	2	
190-199			8	4	2	
200-209			5	4	3	1
210-219		2	11	4	2	1
220-229		3	9	0	2	
230-239		2	2	5	0	
240-249	3	2	6	5	1	1
250-259		2	5	1	1	
260-269	1	1	7	2	0	
270-279		2	4	2	0	
280-289	1	1	7	2	2	2
290-300		4	15	1	5	2
Incomplete	17	31	51	43	8	6
Total	25	63	179	105	45	16
Median	Inc.	300	249.2	247.0	200.5	
Median age	5.71	6.51	7.57	8.43	9.51	

Here again we find that the number of moves does not serve to differentiate between the ages above six.

Time differences appear in these data, but a complication arises in the fact that the 7 and 8 year old children have practically the same median performance. On the other hand, it is particularly noteworthy that two-thirds of the five year olds tested, one-half of the 6 year olds, and one-third of the 7 and 8 year olds *fail* to complete the test within the time limit of 5 minutes.

TENTATIVE STANDARDS FOR FORM BOARD 1C*

In order to give some basis for classification, the following scheme is proposed.

Failure after 5 minutes—Performance Age of 6 or less.

Completed in 4 to 5 minutes—Performance Age of 7-8 years.

Completed in 3 to 4 minutes—Performance Age of 9 or above.

Add one year to "P.A." for completion of test in from 12 to 15 moves.

---

*Much smaller time differences than are here recognized may be significant as will be seen from the following results of Dr. R. E. Leaming with this test. The test was given to 600 fifteen year old boys and girls, divided into three groups:—the so-called "job hunting" group—the least successful of working children, "who could not get a job for themselves, could not hold a job once they secured it, or did not know where to look for a job;" the "job holding" group, 100 boys and 100 girls, then in continuation school and considered as representing the better grade of working children; and a high school group of 100 boys and 100 girls. Group IV is another high school group of 300 girls, and Group V a combined group—400 girls, Group IV, and girls of Group III. "Min." in the table indicates lowest score in seconds, and "Max." the highest. 20% Med. and 80% are the 20, 50, and 80 percentiles.

10. Dearborn (First Trial, Seconds)

	Boys					Girls				
	Min.	20%	Med.	80%	Max.	Min.	20%	Med.	80%	Max.
Group I	45	85	160	320	510 + 5	43	133	228	440	562 + 8
Group II	24	63	118	579	579 + 3	36	105	157	270	532 + 2
Group III	42	70	122	220	495 + 1	43	73	166	251	420 + 2
Group IV						53	100	180	320	580 + 12
Group V						43	94	175	300	580 + 14

11. Dearborn (Second Trial, Seconds)

Group I	20	45	77	127	442	33	75	110	197	322 + 2
Group II	26	38	63	156	315	25	49	70	147	399
Group III	20	40	55	105	190	24	40	70	120	395
Group IV						32	65	107	203	555 + 5
Group V						24	57	97	180	555 + 5

FORM BOARD 2

The Reconstruction puzzle, Form Board 2, has not been used since the early experiments showed its difficulty for the younger children. Table 7 shows the results obtained by testing 30 children. There is in general a decrease in the average time as the age of the children in the groups increases, though there is an exception to this in the 12-13 group. The number of cases is so small that it is impracticable to determine even tentative standards.

TABLE 7

Form Board No. 2—Distributions of Time by Ages

Age	Cases	1	2	3	4	5	6	7	8	Av.
4-7	3	162.0	118.3	69.0	90.0	147.0	53.5	11.5	68.0	97.8
8-9	5	103.1	117.0	94.2	136.2	100.7	35.0	33.3	63.0	95.9
10-11	10	82.5	53.0	42.4	54.7	55.3	46.2	37.3	35.4	50.8
12-13	8	40.1	92.1	49.3	61.8	27.2	90.5	73.0	87.5	66.3
14-16	4	82.0	68.0	24.5	55.0	29.5	32.7	37.7	15.2	33.1
All	30	86.1	81.9	53.5	72.6	54.6	57.8	45.6	54.0	63.8

FORM BOARD 3

The distributions of the results obtained on Form Board 3 are shown in Tables 8 to 15 inclusive. Separate distributions are shown for moves and time on each problem, and in addition distributions showing the total time and total number of moves on all three problems are presented. In making this time distribution, the time of all incomplete problems was counted as 300 seconds. In order to compute the median number of moves, the incomplete records were counted as equal to the maximum number of moves found in the completed records. This procedure is, except in rare cases, justified, since the number of moves made in the incomplete records is usually above the median of the number of moves made in the completed records.

NOTE:——Those who did not complete a test with the time limit are recorded as failures. The number of such failures is entered in the table following the highest score made within the time limit. For example, 234 + 2 means that there were two failures to complete the test in the time limit of three hundred seconds. Quoted from sections 10 and 11 of "The Witmer Diagnostic Standards," appended to "Tests and Norms for Vocational Guidance at the Fifteen-Year Old Performance Level," by Rebecca E. Leaming, University of Pennsylvania, Philadelphia 1923.

TABLE 8

Form Board 3b—Distributions of Moves by Ages

Moves	Age 6	7	8	9	10	11	12	13
5	3	8	6	1	10	19	12	10
6	2	8	10	5	4	14	12	4
7	1	6	2	3	2	5	8	1
8	0	7	2	5	7	8	4	3
9	0	7	4	1	2	6	3	2
10	1	7	7	3	5	5	0	1
11	0	4	1	0	4	4	2	2
12	1	1	5	2		1	1	1
13	1	2	4	1		1	1	1
14	0	4	2	1		1		2
15	3	3	2	0		1		2
16	0	1	3	0		0		
17	0	4	1	0		0		
18	0	2	2	1		1		
19	0	1		0	1		1	
20	1	2		0				
21	0	2		2				
22	0	0		0				
23	0	1	1	0		1		
24	1	0	1	1				
25	1	1						
26	0							
27	0							
28	1			1				
29								
30		1						
30 plus		1	1		2		1	
Incomplete	2	8	1	2	0	1		
Totals	18	81	55	29	37	68	45	29
Median Score	14.5	10.6	10.5	9.5	8.4	7.2	6.9	7.5
Median Age	6.66	7.53	8.24	9.37	10.21	11.12	12.40	13.36

TABLE 9

Form Board 3b—Distributions of Time by Ages

Seconds	Age 6	7	8	9	10	11	12	13
0-9								
10-19			1			3	2	1
20-29	1	3	1	2	4	6	6	4
30-30	2	4	4	2	6	14	8	8
40-49	0	10	7	3	4	8	7	0
50-50	1	6	6	4	5	5	3	2
60-69	2	7	6	3	5	9	8	2
70-79	0	5	6	0	2	5	3	2
80-89	1	4	4	2	1	2	3	0
90-99	1	3	5	3	2	3	1	3
100-109		4	3	2	0	4		0
110-119		2	3	1	1	1		0
120-129		10	1	2	0	1		4
130-139		2	1	0	2			2
140-149	2	0	0	1	0		1	0
150-159	0	3	0	1	1		2	1
160-169	0	1	1	0	0			
170-179	2	0	2	0	1	1		
180-189	0	0	1	1	1	2		
190-199	1	1	1		0	1		
200-209	0	0	1		1	1		
210-219	2	1						
220-229	1	2						
230-239		0						
240-249		1				1		
250-259		1			1		1	
260-269		1						
270-279								
280-289								
290-300		2						
Incomplete	2	8	1	2		1		
Total	18	81	55	29	37	68	45	29
Median	145	95.0	74.2	82.5	59.0	56.0	49.3	57.5
Median Age	6.66	7.53	8.24	9.37	10.21	11.12	12.40	13.36

TABLE 10

Form Board 8d—Distributions of Moves by Ages

Moves	Age	6	7	8	9	10	11	12	13
9		3	7	3	1	3	5	8	4
10		2	9	6	3	4	8	9	3
11		2	6	1	2	2	8	5	5
12		0	6	7	2	6	10	5	4
13		2	8	8	2	3	6	4	4
14		3	5	9	4	4	4	2	2
15		0	7	4	2	1	6	1	0
16		2	5	2	1	1	0	4	3
17		1	3	1	2	1	3	2	0
18		1	2	2	1	1	2	0	3
19		1	2	2	3	2	4	1	1
20			3	1	1	3	3	1	
21			1	0	1	0	3		
22			3	1	1	0	0		
23			2	1	1	2	0		
24			0	0		2	1		
25			1	1		0		1	
26			1	0		0			
27		1	0	0		1			
28			0	1					
29			1						
30			0						
30 plus			1	4	2	1	3	1	
Incomplete			8	1	0	0	2	1	
Total		18	81	55	29	37	68	45	29
Median		14.0	14.9	14.3	15.3	14.1	13.5	12.1	12.6
Median	Age	6.66	7.53	8.24	9.37	10.21	11.12	12.40	13.36

TABLE 11

Form Board 8d—Distributions of Time by Ages

Seconds	Age 6	7	8	9	10	11	12	13
0-9								
10-19			1	1				
20-29		2	0	0	1	2	6	
30-39	1	4	2	2	4	3	5	6
40-49	0	6	3	1	5	6	5	5
50-59	0	6	7	7	2	10	3	1
60-69	3	8	5	4	4	12	6	8
70-79	4	7	8	3	3	3	1	2
80-89	2	3	4	2	4	5	5	2
90-99	0	4	6	1	2	2	4	1
100-109	1	2	2	0	1	4		2
110-119	1	1	2	0	0	0	2	0
120-129	1	5	2	2	4	3	2	1
130-139	0	3	2	0	1	3		0
140-149	1	1	2	0	1	5		0
150-159	1	2	0	1	0	2		1
160-169	2	2	0	1	0	1		
170-179	1	2	0	0	1	0	2	
180-189		4	1	3	1	1	1	
190-199		1	1		1	1	0	
200-209		0	1			1	0	
210-219		2	1			0	1	
220-229		0	1			0	0	
230-239		3	0			1	0	
240-249		2	2		1		1	
250-259		0		1	1			
260-269		3						
270-279								
280-289								
290-300								
Incomplete		8	2			2	1	
Total	18	81	55	29	37	68	45	29
Median	85.0	102.5	83.8	68.8	78.3	73.3	65.9	63.1
Median Age	6.66	7.53	8.24	9.37	10.21	11.12	12.40	13.36

TABLE 12

Form Board 3c—Distributions of Moves by Ages

Moves	Age 6	7	8	9	10	11	12	13
8			1	1	1	4	5	3
9	1	2	0	1	3	9	5	3
10	1	4	4	3	4	6	5	1
11	1	5	2	6	3	3	5	0
12	0	5	2	3	3	9	6	1
13	1	3	7	1	4	3	1	0
14	1	4	1	1	5	4	3	3
15	0	6	10	1	1	2	2	3
16	3	7	3	1	0	5	1	3
17	0	5	0	1	1	3	1	1
18	1	6	1	1	3	1	3	2
19	0	3	2		0	0	0	1
20	2	3	1		1	0	0	1
21	0	3	2		2	2	1	
22	0	1	2		0	0	1	
23	1	3	4		0	2	0	
24	0	6	0		1	2	0	1
25	0	1	2		1	1	2	
26	0	0	1	2	0	1		
27	2	2	1	0	0	1		
28	1	0	2	1	1	1		
29	0	0		0	1	1		
30	0	1		3	0	1		
30 plus	2	2	4	2	2	2	3	4
Incomplete	1	9	3	1	0	5	1	2
Total	18	81	55	29	37	68	45	29
Median	19	17.9	16.2	13.5	14.1	14.0	12.4	16.2
Median Age	6.66	7.53	8.24	9.37	10.21	11.12	12.40	13.36

TABLE 18

Form Board 8c—Distributions of Times by Ages

Seconds. Age	6	7	8	9	10	11	12	13
0-9								
10-19							1	
20-29							1	
30-39				1	2	4	2	1
40-49		3	6	1	1	2	5	1
50-59		4	1	4	1	6	5	1
60-69		7	2	3	6	3	7	4
70-79	3	8	4	1	2	6	2	1
80-89	3	8	4	1	2	6	2	1
90-99	2	7	3	7	4	4	4	0
100-109	0	3	1	0	2	6	1	2
110-119	1	2	5	0	3	3	0	2
120-129	0	6	7	2	1	7	4	4
130-139	1	4	1	0	3	1	0	2
140-149	0	3	3	0	2	2	1	1
150-159	1	2	1	1	1	1	1	1
160-169	1	0	2	0	2	1	2	0
170-179	1	1	0	0	0	2	0	1
180-189	1	3	3	1	2	7	1	
190-199	0	2	1	1	0	0	0	
200-209	1	3		1	0	0	1	
210-219	0	2	0	1	1	1	1	1
220-229	2	0	2	0	0	0	0	1
230-239	0	1	1	0	1	1	0	
240-249	1	1		1		1	1	
250-259	1	0						
260-269		2						1
270-279		2			1		2	1
280-289								
290-300								
Incomplete	1	9	4	2			1	3
Total	18	81	55	29	37	68	45	29
Median	160	121	117	93.6	102.5	97.5	77.5	123.7
Median Age	6.66	7.57	8.24	9.37	10.21	11.12	12.40	13.36

TABLE 14

Form Board No. 3—Distributions of Total Moves by Ages

Moves	Age 6	7	8	9	10	11	12	13
19			1					
20-4		1		2	2	2	5	2
25-9	1	4	5	2	9	18	11	5
30-4	4	10	6	7	6	13	14	6
35-9	1	13	7	4	5	9	6	4
40-4	2	12	13	2	5	8	1	2
45-9	2	13	8	3	3	3	4	3
50-4		6	3			3		3
55-9	2	6	3	1	2	1	2	3
60-4	1	1	2	3		1		
65-9	1	4	3		2	3		
70-4	3	6		3	2	3		
75-9	1	3						1
80-4			1			1	1	
85-9		1	1		1	2		
90-4		1	1	1				
95-9				1		1		
100-4								
105-9								
110-4			1					
115 plus							1	
No. Cases	18	81	55	29	37	68	45	29
Median	47.5	45.2	43.3	39.4	36.5	35.5	33.3	36.9

TABLE 15

Form Board No. 3—Distributions of Total Time by Ages

Seconds	Ages 6	7	8	9	10	11	12	13
60-89							2	
90-119		1		1	2	2	3	
120-149		2	1	2	2	5	10	5
150-179		2	6	6	3	5	4	4
180-209	2	8	8	4	5	14	4	3
210-239	1	5	6	0	5	7	1	2
240-269	0	5	7	5	2	5	7	0
270-299	4	5	4	0	3	6	2	5
300-329	1	8	6	2	4	7	4	2
330-359	0	9	1	1	3	4	1	2
360-389	1	4	2	0	1	6	0	0
390-419	1	2	5	2	1	3	2	1
420-449	0	3	0	2	2		1	0
450-479	0	6	1	0	0		0	1
480-509	2	1	1	1	0		1	2
510-539	3	2	2	0	3	3	1	1
540-569	1	3	0	2	1		0	0
570-599	1	2	1	0			1	0
600-629		2	1	1				1
630-659		1	0					
660-689		4	2			1		
690-719	1	2	0				1	
720 plus		4	1					
Totals	18	81	55	29	37	68	45	29
Median	390	345	268	249	263	246	206	273
Median Age	6.66	7.51	8.24	9.37	10.21	11.12	12.40	13.36

The group to which this test was given is much larger than any of the other groups, and the range is greater. The number at each age, however, is small, and in the 6th, 9th, 10 and 13th years especially the medians are not reliable on this account. Further, it is probable that the six-year-olds are a somewhat superior group, and the thirteen-year-olds are probably inferior since none of the group had left the eighth grade.

TENTATIVE STANDARDS FOR FORM BOARD 8

A "performance age" is first determined in the following *table* according to the total number of moves made in doing the three problems.

P. A.	6	7	8	9	10	11	12	13	14
No. of Moves	48	45	42	39	36	34	32	28-31	22-27

To the "performance age" thus found add 1 year if the total time taken for the three problems is *four* minutes or less. If the performance age is 8 or above on the basis of the number of moves, and if the total time taken is over 5½ minutes, subtract 1 year; if the total time taken is over 6½ minutes, subtract 2 years.

THE TRIANGLE BOARD

The Triangle Board (No. 4) has been given to only about 100 children in the first and second grades. The results are shown in Table 16. Thus only at the age of seven are there sufficient cases to determine a fairly reliable median. However, the following tentative standards are suggested for present use.

Age	6	6½	7	7½	8	8½
Seconds	400	350	320	265	230	205

TABLE 16

Form Board No. 4—Distributions of Time by Ages

Seconds	Age 6	7	8	9
20-39				
40-59				
60-79		1		
80-89		0	1	1
110-119	1	0	0	2
140-159		5	4	
120-139		1	1	
160-179		4	2	
180-199	1	8	5	
200-219		6	4	
220-239	1	4	2	
240-259	1	7	3	1
260-279		5	1	1
280-299		3	0	1
300-319		6	1	
320-339	1	6	1	1
340-359		4	2	
360-379	2	1	2	
380-399		1		
400-419		4		
420-439		0		
440-459	1	0		
460-479		1	1	
480-499	1	1		
500-519		1		
520-539				
540-559				
560-679			1	
580-599	1	2		
600 plus		1	3	2
Total	10	72	34	9
Median	350	260	220	270
Median Age	6.83	7.57	8.20	9.36

THE LINCOLN HOLLOW SQUARE

For the Hollow Square (Tables 17 and 18) only a few results have been obtained. These, however, indicate that the test will be useful in ages below 7, for which purpose it was designed.

TABLE 17

Hollow Square—Distributions of Problems Solved by Ages

Problems Age	3	4	5	6	7	8
8			1	2	3	2
7				5	1	5
6		1	4		2	
5			1		1	1
4		2	1			
3	1	1				
2		1				
1						
0						
Total	1	5	7	7	7	8
Median		4	6	7	7	7
Median Age	3.9	4.75	5.75	6.33	7.5	8.5

TABLE 18

Hollow Square—Distributions of Total Time by Ages*

Seconds	Age 3	4	5	6	7	8
400						
380						
300		1	1			
360	1	1				
340						
320		1				
280		1	1			
260			1		1	
240						1
220				1	1	
200		1	3	1	1	1
180						
160				2	1	3
140				2		3
120			1			
100						
80					2	
60				1	1	
40						
20						
0-19						
Totals	1	5	7	7	7	8
Median	374†	310	216.6	165	170	166.6
Median Age	3.9	4.75	5.75	6.33	7.5	8.5

*In calculating the total time each incompleted problem is given a score of 60 seconds.

†Actual score.

Two methods of scoring have been considered, first counting the problems correctly solved, and second, scoring on the basis of total time necessary for all eight problems. In the latter case the time of each problem not completed within the time limit was recorded as 60 seconds. The data which are at hand indicate that the time score is the most likely to prove useful, and the following tentative norms are suggested.

Age	3	3½	4	4½	5	5½	6	6½	7
Seconds	450	400	350	320	290	240	190	160	130

PICTURE COMPLETION TEST

The scores on the Picture Completion Test (Table 19) suggest tentative standards for the reason that the median scores differentiate only on a two year basis. The medians for the 7th and 8th years are practically the same. This is true also of the 9th and 10th and of the 11th and 12th. The differentiation between these three levels is, however, well marked.

TABLE 19

Picture Completion Test—Distributions of Scores by Ages

Score	Age 6	7	8	9	10	11	12
100							
95-9			1			3	1
90-4			1		1	1	2
85-9		1	1	1	2	1	3
80-4		2	1	4	1	5	2
75-9		1		2	3	2	1
70-4		1	4	2			2
65-9	1		1	2	1	1	1
60-4	1	1	3	1	2		
55-9			1		1	1	
50-4		1					
45-9		1	1				
40-4		1	1				
35-9	1	1	2	1			
30-4	1		2				
25-9	1						
20-4	1						
15-9							
10-14		1					
5-9					1		
0-4			1				
Total	6	11	20	13	12	14	12
Median	35.0	62.5	63.3	76.3	76.6	84.0	85.0
Median Age	6.0	7.58	8.33	9.58	10.17	11.46	12.46

There is a complication in the time scores of this test (Table 20) in that the 6 and 7-year-old children made considerably faster times than the 8-year-olds. Further cases will be necessary to determine whether the 8-year median is too high or the 6 and 7-year medians are too low.

TABLE 20

Picture Completion Test—Distribution of Time by Ages

Minutes	Age 6	7	8	9	10	11	12
15 plus	1	1	3		1		
14.5							
14		1					
13.5			1				
13	1		1				
12-5							
12			2	2			
11.5			1				
11			1	1			
10.5			1				
10		1	1		1		
9.5				1			
9		1					
8.5		1	2		1		1
8		1			1	2	
7.5		1	1	1			
7			1				1
6.5			1				
6	1	1	1	3	1	1	1
5.5	1	1		1			
5	2	1			2	1	1
4.5			1		1	4	1
4			1	1	1	2	1
3.5			1	1	1	3	1
3				1	1		3
2.5		1		1	1	1	
2							2
Total	6	11	20	13	12	14	12
Median	6.0	8.25	10.5	6.25	5.25	4.625	4.0
Median Age	6.0	7.58	8.33	9.58	10.17	11.46	12.46

FURNITURE TEST

Table 21 shows the distribution of scores on the Furniture Test. It is clear that scores do not differentiate above the age of 9.

TABLE 21

Furniture Test—Distributions of Score by Ages

Score	Age 7-8	9	10	11	12	13
15	4	3	5	13	16	8
14		2	3	2	3	3
13	1					1
12						
11	1					
9	1		2	3	2	1
8						
7				1	2	1
6						
5	1		1	2	1	1
4						
3						
2						
1						
0	4	1		1		
Total	12	6	11	22	24	15
Median	10	15	15	15	15	15

In Table 22 are the time distributions on this test. We find a falling off in time with the increase in age, but there are some irregularities. Further testing will be necessary before final standards can be obtained.

TABLE 22

Furniture Test—Distributions of Time by Ages

Minutes	Age 7-8	9	10	11	12	13
15 plus	1	2	1	2	1	3
14.5					1	
14.0				1		
13.5	1					
13.0				1	1	
12.5					1	
12.0						1
11.5	1		1	1	1	
11.0	2		1		1	1
10.5				1	1	1
10.0				1		
9.5						
9.0	1			1		
8.5	1		1	1	2	1
8.0	1				2	
7.5	1			2		
7.0						
6.5		2	1			
6.0				2		
5.5	1			1	1	
5.0			1			1
4.5		2	2	2	2	1
4.0				2	2	3
3.5	1			1	4	
3.0	1		1	2	2	1
2.5			1	1	1	1
2.0			1			
1.5					1	1
1.0						
0.5						
Totals	12	6	11	22	24	15
Median	9.0	6.75	6.0	7.0	5.25	5.25
Median Age	8.19	9.63	10.16	11.42	12.5	13.33

WHAT DO THE FORM BOARDS MEASURE?

The results presented in the foregoing section may have raised in the mind of the reader the questions as to whether performance tests really measure the same thing which we measure with the Binet type of scale and call "general intelligence." There is most certainly a possibility that the two types of tests measure different traits. Thorndike has pointed out that there are three intelligences which everyone possesses, the abstract intelligence, the mechanical and the social intelligence.*

Terman indicates a belief that form boards do not measure intelligence, as he defines it.† In the army examining it was proposed to use the Stenquist Construction test for rating the mental ability of the men who could take the more linguistic examinations, but this had to be abandoned because it was found that feebleminded inmates of institutions frequently made high scores on this test, and the correlation between the scores of unselected men on the Stenquist and examination A were uniformly low.‡

It seems altogether probable that the various types of performance tests differ somewhat as to the traits which they measure. Form boards like No. 3 perhaps measure rather specifically the conception of form and manipulative skill, while those like the Picture Completion and Furniture tests are more general in their nature and require abstract thinking in some degree for their successful completion. It is possible that we shall discover that performance tests of general intelligence must be of this latter type.

Whether or not we find that performance tests are a measure of general intelligence, however, there can be no doubt of their usefulness as a part of the clinical examination. The subject with a low I. Q. who does well in the form board tests can probably be developed into at least a semi-skilled worker, while the possibilities for the education of the feebleminded or borderline case with poor performance test scores are very much more limited. Thus it seems highly desirable to supplement the general intelligence examination with performance tests.

---

*Thorndike, E. L., Intelligence and Its Uses. Harper's Magazine, 1920, Vol. CXL, p. 227. See also his contribution to Symposium on Intelligence and Its Measurement, Journal of Educational Psychology, 1921, Vol. XII, No. 3, p. 124.

†Terman, L. M., Contribution to the Symposium above cited.

‡Psychological Examining the U. S. Army, Memoirs of the National Academy of Sciences, Vol. XV, p. 321.


RARY
ned on
ow